Analysis of Movement of Cellular Oscillators in the Pre-somitic Mesoderm of the Zebrafish Embryo

DISSERTATION

zur Erlangung des akademischen Grades

Doctor rerum naturalium

(Dr. rer. nat.)

vorgelegt
der Fakultät Mathematik und Naturwissenschaften
der Technischen Universität Dresden

von

Bhavna Rajasekaran

geboren am 17 Februar 1983 in Chennai, Indien

Gutachter:

Prof. Dr. Frank Jülicher

Prof. Dr. James Sharpe

Eingereicht am: 30. Oktober 2012

Tag der Verteidigung: 13 Februar 2013

To my family

We are limited by our own abilities.

Contents

Chapter 1

Abstract

During vertebrate embryo development, the body axis is subdivided into repeated structures, called somites. Somites bud off from an un-segmented tissue called the pre-somitic mesoderm (PSM) in a sequential and periodic manner, tightly controlled by an in built molecular clock, called the "segmentation clock". According to current understanding, the clock is comprised of: (i) an autonomous cellular oscillator consisting of an intracellular negative feedback loop of Her genes within the PSM cells, (ii) Delta-ligand and Notch-receptor coupling that facilitates synchronization of oscillators among the PSM cells, (iii) Tissue level waves of gene expression that emerge in the posterior PSM and move coherently towards anterior, leading to global arrest of oscillations in the form of somites. However, the movement of cellular oscillators within the PSM before the formation of somitic furrows, a prominent feature of the tissue as observed through this work has not been experimentally considered as a constituent of the segmentation clock so far. Our work aims to incorporate movement of cellular oscillators in the framework of the segmentation clock. It is well known from theoretical studies that the characteristics of relative motion of oscillators affect their synchronization properties and the patterns of oscillations they form. Particularly, theoretical studies by Uriu *et al.*, PNAS (2010) suggest that cell movements promotes synchronization of genetic oscillations. Here, we established experimental techniques and image analysis tools to attain quantitative insight on (i) diffusion co-efficient of cellular oscillators, (ii) dynamics of a population of oscillators, within the PSM aiming towards concomitant understanding of the relationship between movement and synchronization of cellular oscillators.

In order to quantitatively relate cellular oscillators and their motion within the PSM, I established imaging techniques that enabled visualization of fluorescenctly labeled nuclei as readouts of cell positions in live embryo, and developed dedicated segmentation algorithm and implemented tracking protocol to obtain nuclei positions over time in 3D space. Furthermore, I provide benchmarking techniques in the form of artificial data that validate segmentation algorithm efficacy and, for the first time proposed the use of ‘transgenic embryo chimeras‘ to validate segmentation algorithm performance within the context of *in vivo* live imaging of embryonic tissues. Preliminary analysis of our data suggests that there is relatively high cell mixing in the posterior PSM, within the same spatial zone where synchronous oscillations emerge at maximum speed. Also, there are indications of gradient of cell mixing along the anterior-posterior axis of the embryo. By sampling single cell tracks with the help of nuclei markers, we have also been able to follow *in vivo* protein oscillations at single cell resolution that would allow quantitative characterization of coherence among a population of cellular oscillators over time. Our image analysis work flow allows testing of mutant embryos and perturbation of synchrony dynamics to understand the cause-effect relationship between movement and synchronization properties at cellular resolution. Essentially, through this work, we hope to bridge the time scales of events and cellular level dynamics that leads to highly coordinated tissue level patterns and thereby further our understanding of the segmentation clock mechanism.

Chapter 2

Introduction

During vertebrate embryo development, the body axis is divided into segmented structures, called somites. These structures give rise to the vertebrae, skeleton and the muscle of trunk and tail. Segmentation of the embryo occurs as it grows, in an anterior-to-posterior or head to tail sequence, where somites appear sequentially on the paraxial mesoderm tissue, that is differentiated much earlier during development (Wolpert et al., 2007; Gilbert, 2006). Although, we do not completely understand the mechanism controlling the process, we do know some of the key players that determine somite formation. Studies have proven that a biological clock and wave mechanism based on gene expression network of oscillators within the mesoderm tissue coordinate to give rise to the sequential and periodic somite pattern, harmonized by other signaling pathways- the process termed ‘somitogenesis‘ and the clock referred to as ‘segmentation clock‘. This manner of biological patterning has attracted researchers who are not only biologists, but also physicists, computational biologists and recently even engineers. I will introduce and explain the mechanism by which vertebrate embryos are segmented in the following section.

2.1 Somitogenesis during zebrafish embryo development

During the process of somitogenesis, the body axis of the vertebrate embryo is segmented in an anterior-to-posterior sequence, where groups of cells called somites –precursors of vertebrae, ribs, and some dermis and muscular tissue, appear sequentially from head to tail as the embryo grows (Kimmel et al., 1995; Dequéant and Pourquié, 2008; Wolpert et al., 2007; Pourquié, 2001; Gilbert, 2006). In zebrafish, around 30 somites are formed in the paraxial mesoderm with one new pair of somite formed about every 30 minutes from the posterior unsegmented tissue of the embryo, the presomitic mesoderm (PSM) as shown in Fig. 2.1. Somite formation is hence a highly time-driven process, with a species specific period, ranging from 30 minutes in zebrafish to 6 hours in human (Gomez et al., 2008). In the following, I will give a brief historical account of the most important steps leading to our present understanding of the process of somitogenesis. The first theoretical hypothesis to explain somitogenesis was the 'clock and wavefront model' proposed by Cooke and Zeeman (Cooke, 1975; Cooke and Zeeman, 1976). According to their model, tissue level synchronous oscillations occur in the PSM, which upon reaching the determination front in the anterior, arrest to form somites. The length of the somite formed is given by the product of the arrest wavefront velocity and the period of the clock. The model thus explains how temporal periodicity is converted into spatial patterns (Cooke and Zeeman, 1976; Morelli et al., 2009). However, the model doesn't take into account the cellular and molecular players involved in the process of somite formation.

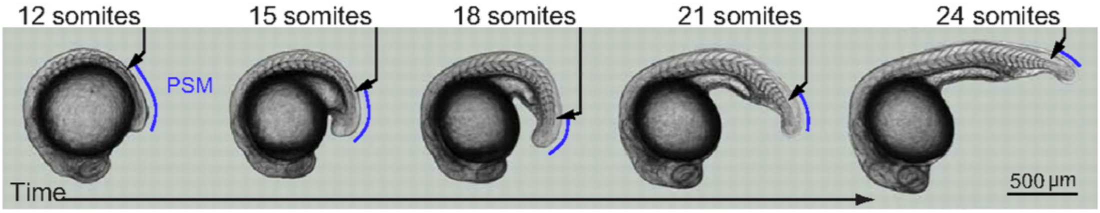


Figure 2.1: Lateral views of embryo with stages during zebrafish embryo somitogenesis. Somites bud-off sequentially on the posterior unsegmented PSM (blue line). courtesy: Oates A C *et al.*, Development, 2012.

The first molecular evidence for a developmental clock linked to segmentation and somitogenesis of the paraxial mesoderm was provided for the chick embryo (Palmeirim et al., 1997). *Palmeirim et. al.* identified and characterized a gene belonging to the *hairy and Enhancer of split* (Hes) family (Davis and Turner, 2001; Kageyama et al., 2007), (*c-hairy1*), where its mRNA exhibited cyclic waves of expression within the paraxial mesoderm alone, whose temporal periodicity corresponds to the formation of one somite (Fig. 2.2 A.). Similar orthologous sets of Hes-related (Her) genes have been identified in the segmentation clocks of other vertebrates (Krol et al., 2011). The cyclic nature of *deltaC*, that codes for Notch ligands, was revealed through in situ expression pattern analysis at different stages of zebrafish somitogenesis. These findings suggested the existence of a segmentation clock in the zebrafish PSM and the crucial function of Notch signaling to synchronize the biochemical oscillators within the PSM of the zebrafish embryos (Jiang et al., 2000). Further investigations on the roles on two genes, *her1* and *her7* revealed that they repress the transcription of their own protein, and their mRNA levels oscillate in synchrony, with a periodicity and phase matching those of *deltaC* (Hirata et al., 2002; Henry et al., 2002; Oates and Ho, 2002; Holley et al., 2000; Lewis, 2003) (Fig. 2.2 B.).

A large genetic screen identified several zebrafish segmentation mutants such as *fused somites (fss), beamter (bea), deadly seven (des), after eight (aei) and mind bomb (mib). After eight (aei) and beamter (bea)* are caused by mutations in *deltaD (deld)* and *deltaC (delc)*, respectively, while *des* is a mutant of the *notch1a* receptor and *mib* is an E3 ubiquitin ligase required for ubiquitination of the Delta ligand, suggesting that in some way or the other, all of these mutants affect the Delta/Notch signaling pathway. The expression pattern of *delc* in all the delta-notch mutant backgrounds was shown to be clearly erratic, as seen in Fig. 2.2 C, further supporting the requirement of the Delta/Notch signaling to keep the oscillations of neighboring PSM cells synchronized (van Eeden et al., 1996; Jiang et al., 1996; Jülich et al., 2005; Jiang et al., 2000; Itoh et al., 2003).

Disruption of the signal leads to salt and pepper like patterns of oscillating gene expressions. Embryos treated with DAPT, a γ - secretase inhibitor that blocks the cleavage of the Notch-intra cellular domain (NICD), that normally regulates the transcription of *her* genes, leads to disruptive segmental boundaries, revealed by high resolution in situ hybridization as a read out. As a consequence of loosing signaling, local coordination among the cells is disrupted, although the cells would still presumably continue to oscillate in the PSM. Wash-out of DAPT from the embryo lead to complete recovery of segmental boundaries as shown in Fig. 2.3. These experiments strongly suggested that the cell autonomous oscillations are

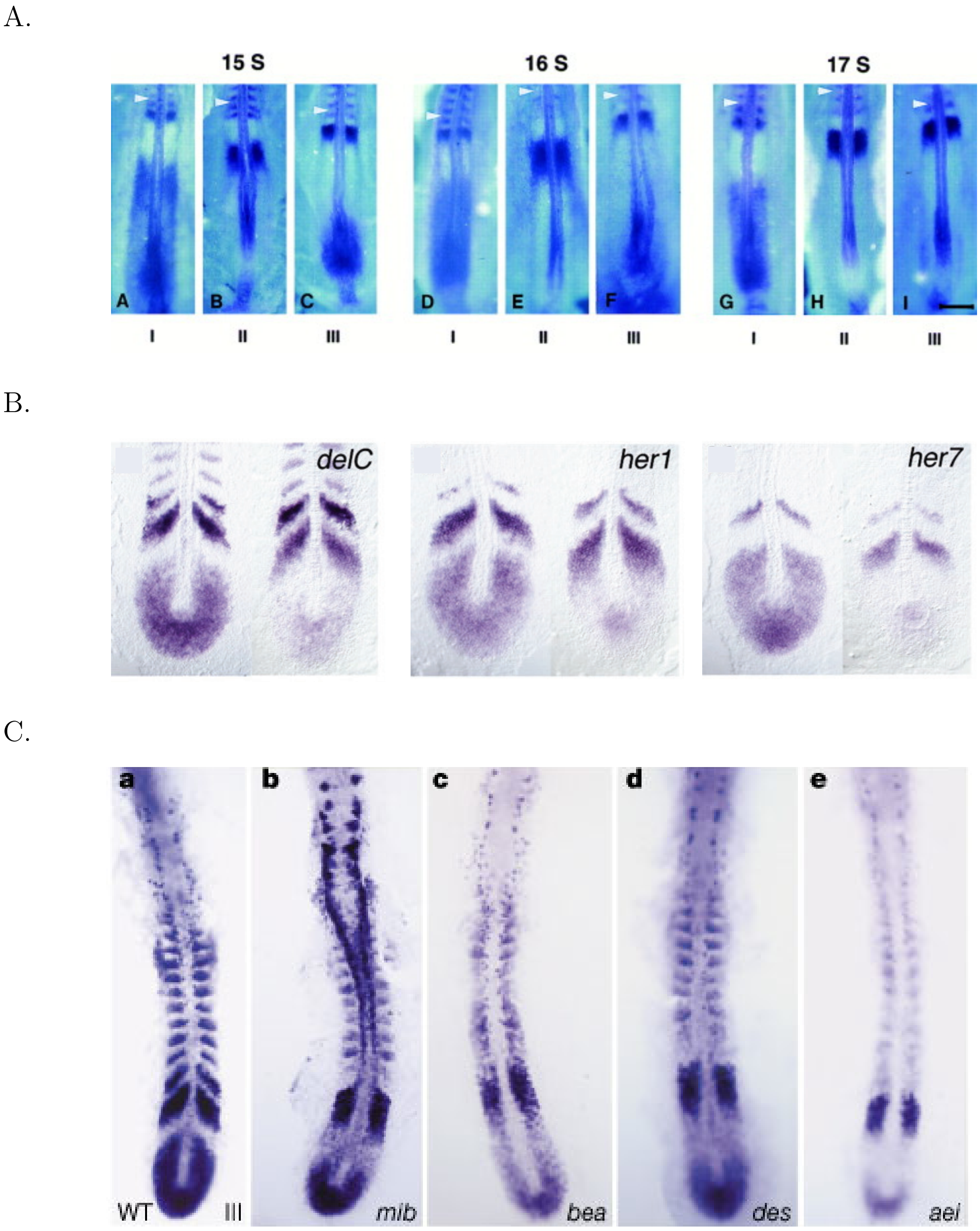


Figure 2.2: Dorsal views A. In situ hybridization with *c-hairy1* mRNA expression in the PSM of chick embryo provides first molecular evidence for oscillatory gene expression. *c-hairy1* expression patterns in embryos aged, 15 (A, B, and C), 16 (D, E, and F), and 17 (G, H, and I) somites. I, II, III for each somite stage indicates the periodic phase of oscillations in the PSM; adapted from: Palmeirim *et al.*, Cell, 1997. B. In situ expression of cyclic genes in the zebrafish PSM at 14 hpf (10 somites) after flat mounting for *delC, her1, her7*; adapted from: Oates *et al.*, Development,2002. C. *papc, (paraxial protocadherin)*, marks cells corresponding to the anterior part of a somite, (Yamamoto et al., 1998)) and *delc* expression patterns in wild type (wt) and mutants (*mib, bea, des, aei*), at 10-somite stage. *delc* expression pattern in all mutants is abnormal compared to wt, with no sign of the organized oscillation as seen in the wt, indicating individual cells continue to oscillate in the mutants with an amplitude appropriate to the region of the PSM in which they lie, but that local cell-cell synchrony is lost, creating a random mixture of cells in different phases of their individual oscillation cycles. Taken from: Jiang *et al.,* Nature, 2000.

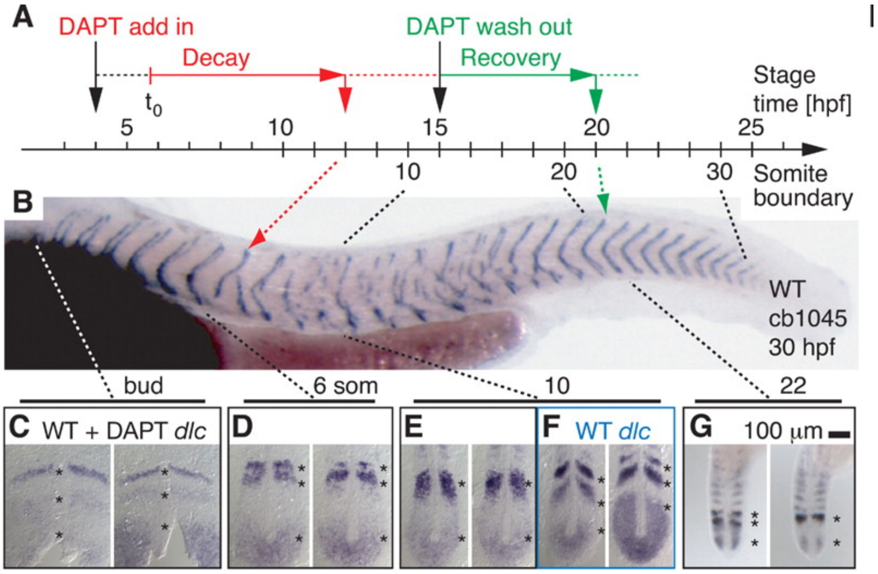


Figure 2.3: Lateral view, 36 hpf. Delta-Notch coupling is sufficient for self-organized resynchronization of the segmentation clock and rescue of morphological segmentation defects. A Timeline of DAPT pulse-chase experiment. B Six anterior segments formed correctly followed by approximately 15 disrupted segments as an effect of DAPT, followed by normally shaped segments posterior to 21st segment as an effect of wash-out of DAPT. C to G Representative *dlc* PSM expression patterns, time points marked referring to B. Two representative embryos showing transition from mildly affected cyclic expression stripes C, degrading gradually D to typical salt-and-pepper pattern indicating desynchronization among cells E compared to expected wt pattern F, and return of normal, symmetrical cyclic gene expression pattern G. Asterisks mark *dlc* cyclic stripes (C, F and G) and disordered expression domains (D and E). Taken from: Reidel-Kruse IH *et al.*, Science, 2007.

generated by *her1* and *her7*, and that the Delta-Notch signaling is required to maintain synchrony in the segmentation clock by coupling the cellular oscillators. It is also noted from these studies that the onset of segmentation defects and recovery from the defects occurs after certain 'time delay' (Riedel-Kruse et al., 2007; Jiang et al., 2000; Horikawa et al., 2006; Ozbudak and Lewis, 2008; Giudicelli et al., 2007).

Further, other signaling pathways like Wnt, fibroblast growth factor (FGF) and retinoic acid have been identified to have a gradient like distribution, along the anterior-posterior axis of the PSM, orchestrating tissue level spatio-temporal aspects of somitogenesis (Dequéant et al., 2006; Dubrulle et al., 2001; Pourquié, 2003; del Corral et al., 2003; Dequéant and Pourquié, 2008; Niwa et al., 2011;

Pourquié, 2011; Vilhais-Neto et al., 2010; Oates et al., 2012) as seen in Fig. 2.4.

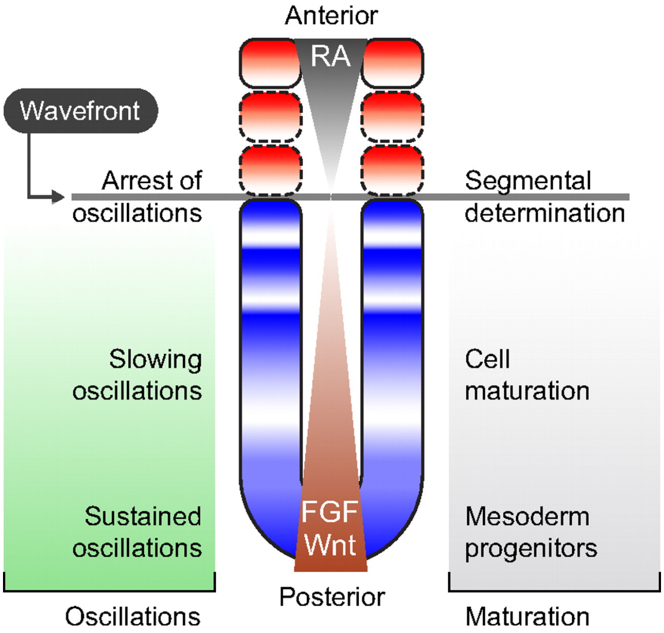


Figure 2.4: Pictorial representation of the dorsal view showing global gradients of fibroblast growth factor (FGF) and Wnt (brown) form a gradient on the tissue from posterior to anterior, while retinoic acid (RA) gradient (gray) expands from the recently formed somites. Sustained high-frequency oscillations observed in the posterior region of the tissue that gradually slow down as they approach the wavefront of arrest (horizontal line). Left panel shows the different phases cellular oscillators undergo as they traverse the tissue. Right panel highlights maturation program of pre-somitic mesoderm cells, running in parallel to segment length specification controlled by the clock. Taken from: Oates A C *et al.*, Development, 2012.

Progress towards understanding the structure and function of the clock mechanism in the last decade manifest combinatorial approaches based on experimental observations, computational analysis and theoretical models (Tomlin and Axelrod, 2007; Oates et al., 2009; Roellig et al., 2011; Lewis, 2008). These findings improved our understanding of the function and dynamics of the segmentation clock, particularly the interplay between single cell oscillator to collective coherent dynamics among a population of cells, leading to tissue level patterns in the form of somitic boundaries (Pourquié, 2011; Oates et al., 2012; Morelli et al., 2012).

Keeping with the pace of appreciating *in vivo* clock dynamics, the idea of simultaneous direct observation of rapidly changing gene expression seemed to quickly build up among researchers studying zebrafish and other vertebrate species. Designing a cyclic gene reporter can be a challenging task since it can be tricky to find the right balance between accuracy of expression, which depends on the instability of reporter proteins, and the sensitivity of available imaging methods/set-ups, which must detect the expected transient signals (Soroldoni and Oates, 2011). The first live reporter of cyclic oscillations using a destabilized Luciferase cassette reporter at the mRNA and protein level, was observed in the mouse embryo by *Masamizu et. el* (Masamizu et al., 2006). A while later, Aulehla and co-workers provided the first fluorescent transgene cassette reporter of oscillating gene expression that could be imaged at the tissue level (Aulehla et al., 2008). Mathematical models of the segmentation clock predict that the oscillatory expression is regulated by negative feedback and 'delays' from transcription to protein expression and therefore negative auto-regulation must be sufficiently long for sustained oscillations (Lewis, 2003; Monk, 2003; Novak and Tyson, 2008). In another study on the intron length in genes, it has been shown that in negative feedback systems, intron length in genes can be significantly instrumental in causing transcription delays and thereby contribute to timing mechanisms of gene expression (Swinburne and Silver, 2008; Swinburne et al., 2008). The function of introns in the timing of gene expression in segmentation clock network to produce sustained oscillations was demonstrated for *Hes7* in mice by *Takashima* and co-workers (Takashima et al., 2011; Oswald and Oates., 2011).

In our laboratory, the PhD work of Dr. Soroldoni has produced high-fidelity live reporters, a technically challenging task, that recapitulates the endogenous Her1 and Her7 oscillating protein expressions (Soroldoni, 2010). Comparative statistical significance tests conducted on live reporters and wt embryos, indicated a consistency between *in vivo* protein oscillations and the period of somite boundary formation that also overlap closely with the wt embryos period (Soroldoni, 2010; Schröter et al., 2008). These transgenic reporters now allow us to follow tissue level oscillating protein expression as seen in Fig. 2.5. High magnification imaging also allows us to study oscillations of these reporters at cellular resolution.

Although several components of the clock have been identified, direct visualization of cellular oscillators is essential for bridging between the cellular level events

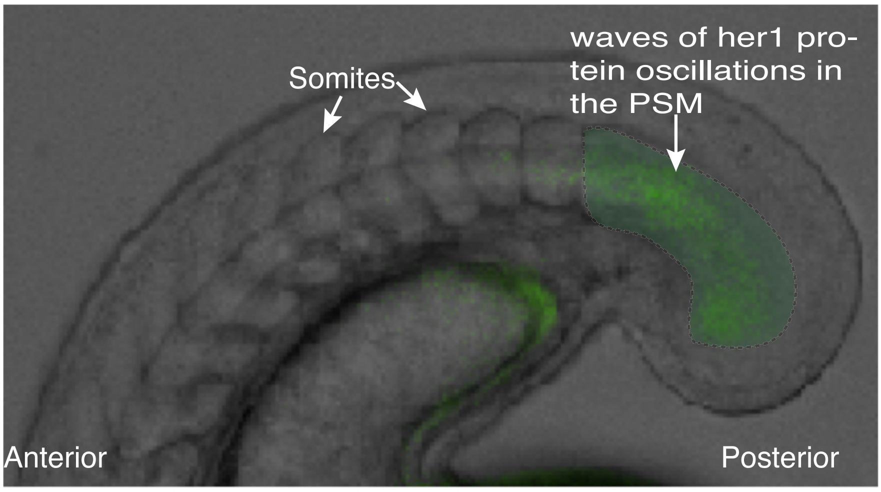


Figure 2.5: Lateral view of the embryo. Tissue level protein oscillations of *her1* - yfp, that occur autonomously in PSM cells. PSM cells are synchronized with adjacent cells, a wave-like gene expression starts at the posterior PSM and moves towards the anterior, where it slows down and eventually stops, where somites are formed. Movie snapshot and transgenic line: Daniele Soroldoni

leading to tissue level patterns derived from a highly coordinated temporal program. Moreover, none of the studies so far considered the role of movement of cellular oscillators within the PSM in the context of the segmentation clock. In the following sections, I will introduce the current state of the art for the segmentation clock. My approach in understanding the role of movement of oscillators is predominantly based on developing tools to obtain quantitative experimental data of single cell trajectories in the PSM. I do this, in order to gain cellular level insight on the clock dynamics. Further, existing and relevant theoretical models are improvised to extract meaningful properties of the system at cellular resolution. Therefore, in subsequent sections, I will introduce the theoretical concepts required for understanding the clock mechanism. Subsequently, I will put forth the current challenges in image acquisition techniques and image processing approaches to acquire 3D positions of cells over time in live embryonic tissues.

2.2 Segmentation clock: the three-tier model

The current understanding of the segmentation clock can be conceptually subdivided into a three-tier model, wherein all the tiers must orchestrate for proper functioning of the clock (Oates et al., 2012; Oates and Ho, 2002; Holley et al., 2002, 2000; Lewis, 2003; Lewis et al., 2009).

Single cell autonomous oscillators: Each cell within the PSM has a protein (Her1 and Her7), which represses the transcription of its own gene (*her1* and *her7*). Periodic changes in concentration of gene and protein via a negative feedback loop, makes the cell an autonomous oscillator, also exemplified in Fig. 2.6 B.

Delta-Notch coupling synchronizes the autonomous oscillators: The mRNA levels of these genes synchronously increase and decrease in a population of cells in the PSM. This is due to the fact that cells in the PSM are coordinated via the Delta/Notch signaling pathway. Notch and Delta are cell-surface receptor and ligand respectively. A delta ligand of one cell, say A, binds to the Notch receptor of its neighboring cell, say B, to activate cleavage of the Notch at the membrane that consequently allows release of the Notch intracellular domain (NICD) in cell B, which migrates to the nucleus and regulates the transcription of *her1* and *her7* genes. The Notch extracellular domain (NECD) along with the Delta is endocytosed in cell B. This causes cells to express similar identities and thereby inter-cellular communication among cells is achieved. Thus, local synchronization is achieved via inter-cellular coupling as shown in Fig. 2.6 C.

Tissue level global control and arrest of oscillations: At the tissue-level, the segmentation clock generates patterns of traveling waves of gene expression, that sweep across the PSM tissue from posterior to anterior. The traveling waves of oscillations are sharply arrested in the anterior PSM in the form of somites, also seen in Fig. 2.4.

Thus the segmentation clock is a multi-scale rhythmic pattern generator, coordinated from a single cell oscillator to coupling of phases of oscillators that integrate to give periodic tissue level patterns (Oates et al., 2012; Pourquié, 2011).

2.3 Moving cellular oscillators of the PSM

Cell movements start very early during embryogenesis and is a fundamental aspect that carves the embryonic body form. Multiple morphogenetic movements occur during development driven by variety of cell behaviors such as cellular rearrangements, cell migrations, coordinated shape changes and region specific

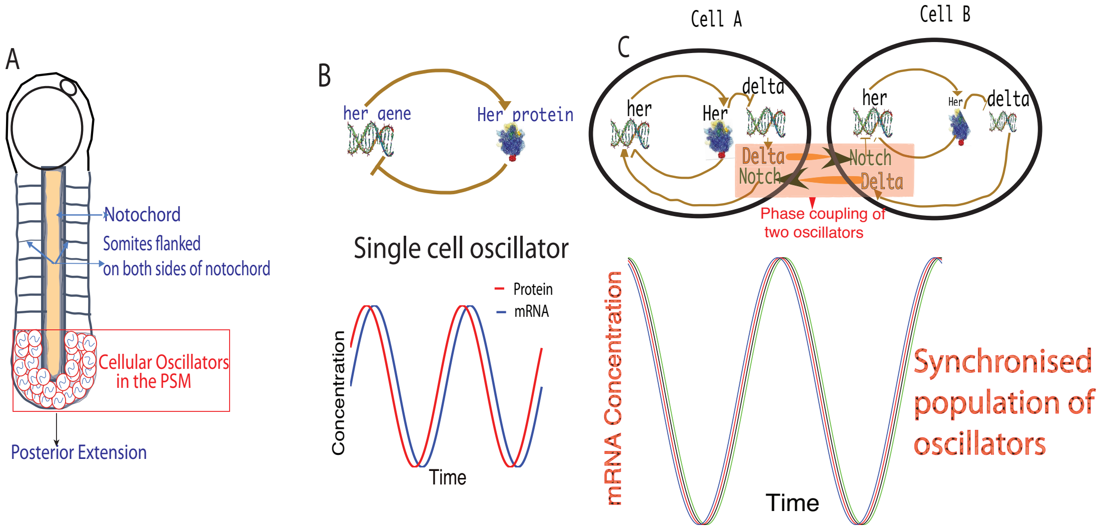


Figure 2.6: A Dorsal view cartoon of the embryo during somitogenesis. Somites flanked on both sides of the notochord. PSM is an assembly of cellular oscillators with a number of cyclic genes expressed in the tissue. B Single cell oscillator via an intracellular negative feedback loop within the nucleus (by Her gene and protein) (top). Oscillations of mRNA and protein concentrations in the single cell oscillator (bottom). C Autonomous Her gene oscillators are synchronized by Delta-Notch coupling (top). Synchronized oscillations of mRNA for a population of coupled oscillators (bottom).

cell movements (Wolpert et al., 2007; Gilbert, 2006). However, cell movement patterns during somitogenesis in the zebrafish have not been elucidated in the context of the segmentation clock. In this study, we aim to incorporate the role of cell movements in the PSM and how their properties affect the synchronization of oscillators. When we look at the cell movements in the PSM over a span of several hours using a Histone H3-gfp transgenic line on a confocal microscope, we perceive continuous dynamic re-arrangment of cells throughout PSM prior to formation of somitic furrows, indicated by white arrows in Fig. 2.7. Maximum cell mixing is observed within the same spatial posterior PSM, where sustained synchronized oscillations occur at high frequency as shown earlier in Fig. 2.4. In this work, we aim to investigate the cause-effect relationship between movement of cellular oscillators and their synchronization properties during the course of somite formation.

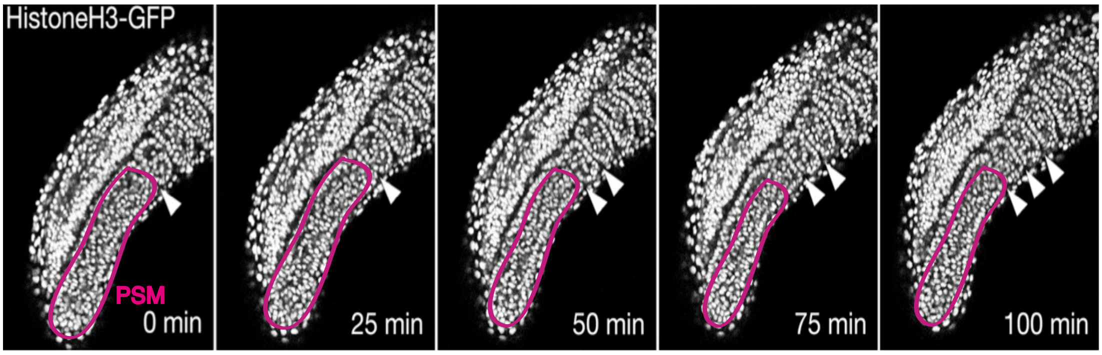


Figure 2.7: Real-time imaging of cell nuclei in zebrafish embryos. Panels are taken from a confocal time-lapse movie (25x objective) and show the positions of nuclei in the tail during somitogenesis; newly forming somite boundaries are marked with white arrowheads. Continuous movement of cells observed throughout out PSM (indicated in pink here). Note that PSM shrinks in length along anterior-posterior axis and decreases in depth along the dorsal-ventral axis as somite formation occurs during development. Snapshots of movie from Dr. Andrew C Oates.

2.4 Relationship between movement and synchrony?

Experimental and theoretical studies on the segmentation clock so far have emphasized the role of coupling of cellular oscillators as an essential criterion for proper functioning of the clock. Coupling or phase synchronization of neighboring oscillators is achieved via interactions between membrane proteins such as delta ligands and notch receptors. A classical experiment shown in Fig. 2.3 further supports that the delta-notch coupling by itself is sufficient to maintain synchrony among a population of oscillators. Other signaling pathways like FGF and Wnt form a gradient along the antrior-posterior axis, however, cell-cell communication alone can be achieved by the delta ligand and notch receptor signaling between cell membranes of neighboring cells, essential for synchronisation. However, delta-notch coupling alone may not be able to attain sustained global level synchronization among a population of cellular oscillators. This is owed to the fact that locally coupled oscillators have strong tendency to form spatial strutters that would prevent the oscillators to achieve global synchronization (Peruani et al., 2010; Uriu et al., 2012). Theoretical studies have thus shown that local spatial structures formation may be overcome by the continuos exchange of oscillator positions to achieve a global synchronous state. Therefore, continuous change of position of oscillators may prevent the formation of local synchronization and rather aid a global synchrony state. These studies suggest that the characteristics of relative motion of oscillators affect their synchronization properties and the patterns of oscillations they form(Ishimatsu et al., 2007; Mehta and Gregor, 2010; Tinsley et al., 2009; Riedel-Kruse et al., 2007; Peruani et al., 2010; Uriu et al., 2010, 2012). Cell movements within the PSM play a key role owed to morphogenesis during embryo development demarcating somitic furrows (Delfini et al., 2005; Bénazéraf et al., 2010) sequentially on the PSM. PSM cells exhibit oscillatory behavior and synchronized oscillations occur in the same spatial zone (posterior PSM) of the tissue where rigorous cell mixing has been observed, one can ask, whether the movement of cells in the PSM has any contribution to the segmentation clock? Some studies have interpreted cell movements as a source of noise that would hamper the oscillatory machinery in its synchronization (Mara and Holley, 2007), however, the theoretical studies prove it different.

Therefore, we ask the following question in the zebrafish embryo PSM:

"What happens to the coupling between cells while the cells move continuously over time within the PSM?"

Or how does the movement of cellular oscillators affect the synchronization properties? To obtain answers, we combined high resolution imaging techniques with quantitative image analysis tools to obtain cellular tracks in order to read *in vivo* oscillator phases. In the next section, I will briefly go through stages of zebrafish embryo development, particularly convergence and extension mechanisms adopted during development and the essential role of cell movements that contributes to this process.

2.5 Cell movements during embryonic ‘convergence and extension‘

Zebrafish are teleosts and form an excellent research model organism since the zebrafish embryos develop externally and are transparent until about 24 hours post fertilization (hpf). The zebrafish developmental stages were first described by *Kimmel et al.,* (Kimmel et al., 1995) in terms of hours post fertilization (hpf) at standard temperature of 28 °C, schematically narrated in Fig. 2.8.

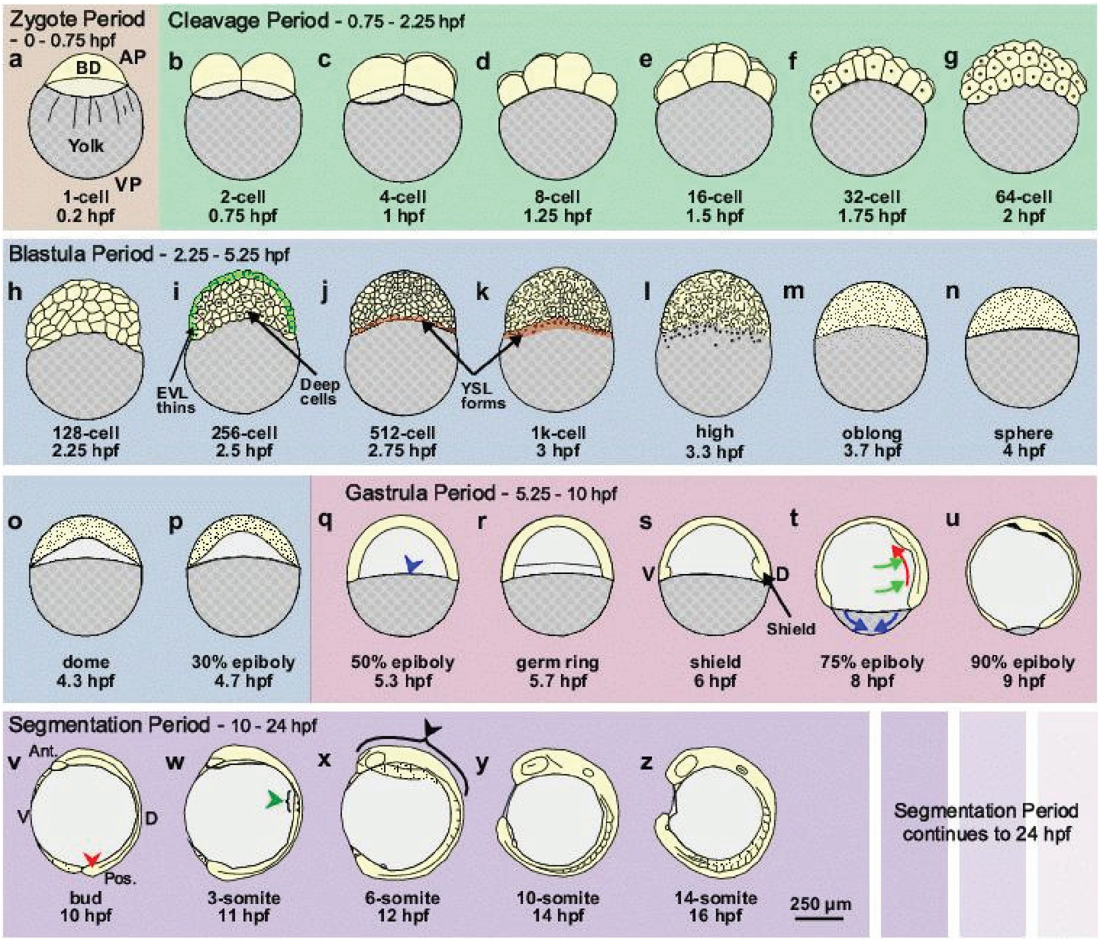


Figure 2.8: Pictorial representation of zebrafish development from the Zygote Period to the mid-Segmentation Period. Following the brief Zygote Period (panel a), when the embryo is at the single-cell stage, the Cleavage Period (panels b–g) runs from 2-cell to 64-cell stage (i.e., 0.75 hpf to 2.25 hpf). The Blastula Period (panels h-p) follows the Cleavage Period, and runs from the 128-cell stage to the 50% epiboly stage (i.e., from 2.25 hpf to 5.3 hpf). The formation of the enveloping layer (EVL) and yolk syncytial layer (YSL) are indicated in panels i, and j and k, respectively. The Gastrula Period (panels q-u) then runs from the end of the Blastula Period at 50% epiboly (i.e., 5.3 hpf) through to bud stage at 10 hpf, after which the SP begins. In this schematic, we have just shown the early stages of the Segmentation Period (panels v-z), i.e., from bud stage (at 10 hpf) to the 14-somite stage (at 16 hpf). AP, VP, Ant. Pos. and hpf are animal pole, vegetal pole, anterior, posterior and hours post fertilization, respectively. BD (in panel a) indicates the blastodisc, at the 1-cell stage, which develops into the blastoderm during subsequent stages of development. In the Gastrula Period, the leading edge of the blastoderm is indicated by a blue arrowhead in panel q. The dorso-ventral axis is first visible morphologically at Shield stage (i.e., 6 hpf; panel s). The simultaneous cell movements of epiboly (blue arrows), convergence (green arrows) and extension (red arrows) that occur during the Gastrula Period are shown in panel t. The tail bud, somites and brain anlage are indicated by red, green and black arrowheads, respectively in panels v, w and x. Schematics modified from Kimmel et al.,. Taken from: *Webb et. al.,* (Webb and Miller, 2007)

The final vertebrate body plan is extremely complex, generated from a single fertilized egg. The fertilized egg, referred to as zygote undergoes several early mitotic cell divisions. The cell - divides, differentiates, grows, moves, produces and senses molecular signals that influence neighboring cell activities, generate patterns. Hence the rules of activities are very well-defined from the beginning for embryo development within a species. Besides, some of the early events are even conserved across species (Alberts et al., 2002). Cell movements start very early during zebrafish development and in fact it has been shown through lineage tracing and fate mapping experiments that cell commitment is set already before the onset of gastrulation (Ho and Kimmel, 1993; Kimmel et al., 1990).

Understanding early morphogenetic processes in developmental biology has been of great interest for several decades (Spemann, 1938; Waddington, 1956). Gastrulation, (refer to Fig. 2.8), a fundamental process of embryogenesis, shapes the early embryo with three germ layers - endoderm, mesoderm and ectoderm. Vertebrate gastrulation involves four morphogenetic movements: mesoderm internalization, epiboly, convergence and extension. During gastrulation, cells demonstrate a repertoire of morphogenetic movements coordinated with fate inductions to sculpt the embryonic body. Epiboly leads to expansion of tissue, internalization encompasses movements of mesodermal and endodermal precursors from the blastula surface beneath the ectodermal layer. Convergence movements narrow embryonic tissue mediolaterally, whereas extension movements elongate them from head to tail (Spemann, 1938; Warga and Kimmel, 1990; Solnica-Krezel, 2005; Keller et al., 2000; Wallingford et al., 2002; Myers et al., 2002).

Convergence movements give rise to directed migration of cells in the lateral regions of the fish embryo toward the dorsal midline without any cellular rearrangements, while the anteroposterior extension encompasses the intercalation of cells specifically along the mediolateral axis via cellular rearrangements (Wallingford et al., 2002; Solnica-Krezel, 2005). A large scale genetic screen, carried out few years ago identified specific mutations affecting cell fates and cellular arrangements during gastrulation of zebrafish embryo development (Solnica-Krezel et al., 1996; Stemple et al., 1996; Sepich et al., 2000).

Two of the mutants described in the screen were reported to have reduced convergence and extension movements, namely the heparan sulfate proteoglycan

Knypek/Glypican4 (Kny) (Topczewski et al., 2001) and Trilobite (Tri)/Strabismus (Stbm)/Van Gogh-like 2 (Vangl2) (Jessen and Solnica-Krezel, 2004; Jessen et al., 2002). *Kny* and *tri* are components of the non-canonical Wnt signaling pathway, also termed as Planar Cell Polarity (PCP) pathway and they control the medio-lateral elongation and alignment essential for the planar cell behaviors that drive convergence and extension processes (Myers et al., 2002; Yin et al., 2000; Sepich et al., 2000). Through quantitative particle image velocimetry analysis (Raffel et al., 1998) of patterns in the PSM and manual cell tracking, it was revealed that the medial PSM cell population in *kny* and *tri* had fewer cell rearrangements and reduced cell migration speeds compared to their wt counterparts, thus supporting the crucial contribution of medial intercalation for convergence and extension movements (Yin et al., 2000; Yin and Solnica-Krezel, 2007; Solnica-Krezel et al., 1996; Keller et al., 2000). As a result, the somites in *kny* and *tri* mutant embryos are substantially wider in their medio-lateral dimension and shorter in their anterior-posterior dimension as seen in figure Fig. 2.9 B.(Solnica-Krezel et al., 1996; Henry et al., 2000). Interestingly, it has been reported by Dr. Röllig in our laboratory, that the period of somitogenesis (Schröter et al., 2008) in the mutants is same as seen in wt embryos and further established a novel role of the PCP pathway in regulation of the cellular morphology during somite boundary formation (Roellig and Oates, 2012). In this thesis, the convergence and extension mutants form our potential candidates for PSM cell movement analysis study, that have slower cell movement compared to wild-type (WT). We also consider to investigate cell movement rates in *aei*, a *deld* mutant, that affects delta - notch coupling between oscillators, in order to understand whether cellular mobility is affected by delta-notch signaling.

2.6 Theoretical studies of the segmentation clock

Mathematical models may serve to provide plausible prototypes, to explain observed phenomena and provide a theoretical framework for further experimental investigations. Theoretical interest in somitogenesis started as early as in the 70's, much before molecular biologists found experimental evidence for a molecular biological clock within the PSM. Christopher Zeeman and Jonathan Cooke

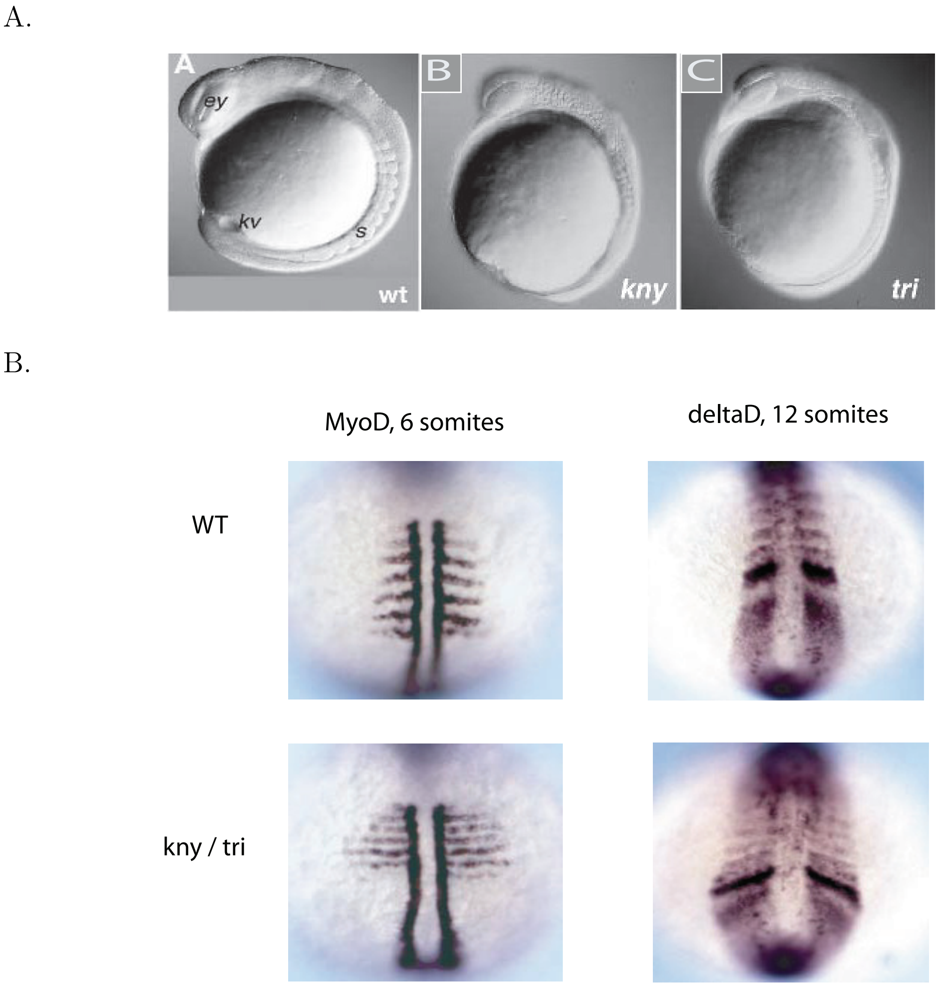


Figure 2.9: Anterior is to the top and dorsal towards the right. A. Lateral views. Nomarski images of live embryos at 10 somite stage (13.25 hpf). Wild type (wt); Mutants: Knypek (*kny*) and Trilobite(*tri*). Scale bar, 0.1 mm. Adapted from: L Solnica-Krezel *et al.*,1996. B. Dorsal views. Top panel wt and bottom panel *kny/ tri* mutant from a cross of two heterozygous adults. MyoD expression indicates early muscle differentiation in both cases at 6 somite stage. DeltaD expression seen at 12 somite stage. Narrow somites are formed in kny and tri embryos compared to wt. Adapted from: Henry *et al.*, 2000.

postulated the *clock and wavefront* model to explain the periodic somite formation during vertebrate morphogenesis (Cooke and Zeeman, 1976), while the existence of an oscillatory gene network responsible for such a sequential and periodic patterning was discovered much later only in the 90's (Palmeirim et al., 1997).

The pioneering work of Turing in 1952 introduced the concept of 'morphogen'and coupled reaction-diffusion systems that could potentially give rise to various types of biological patterns (Turing, 1952). Turing's conceptual idea that mutually-interacting components of a system could spontaneously generate patterns of various forms depending on the dynamics of the reacting components and wavelength of pattern has been widely used to explain pattern formation in various biological systems (Kondo and Miura, 2010). Following this theory, Meinhardt proposed activator-inhibitor systems which further included non-linearities and considered local and global stability criteria to generate spontaneous patterns and structures in tissues (Meinhardt, 1982). In another famous work by Wolpert, it has been suggested that cells process positional information by combining positional identity and a signaling gradient to form patterns during development (Wolpert, 1969, 1994). Since then, considerable experimental evidence has proved the existence of morphogen gradients responsible for cell fate specifications and patterning during development and has inspired theoretical modeling for unraveling the dynamics of biological systems (Gurdon and Bourillot, 2001; Oates et al., 2009; Bollenbach et al., 2005; Dubrulle and Pourquié, 2004; Ashe and Briscoe, 2006; Reeves et al., 2006; Delfini et al., 2005; Bénazéraf et al., 2010; Wartlick et al., 2011; Morelli et al., 2009). Besides, several studies have illustrated the importance of biochemical and biophysical rhythms in various systems ranging from calcium oscillations to circadian rhythms that find their roots in regulatory mechanisms that control the dynamics of living systems (Goodwin, 1963; Winfree, 1980; Goldbeter, 1996), much earlier than the discovery of the segmentation clock (Kruse and Jülicher, 2005). Our present understanding of the molecular clock and gene regulatory network linked to somitogenesis has advanced considerably through quantitative experiments complemented by computational tools and mathematical models.

Soon after the discovery of an oscillatory gene, *c-hairy1* within the PSM (Palmeirim et al., 1997), the first mathematical model for oscillatory expression of the Hes1 protein, a basic helix-loop-helix (bHLH) factor in mouse PSM was presented by Hirata and co-workers (Hirata et al., 2002). They experimentally established that the bHLH Hes1 protein self-represses the transcription factor by inhibiting hes1 mRNA production and show that sustained oscillations depend upon negative auto-regulation of *hes1* transcription and ubiquitin-proteasome-mediated degradation of the Hes1 protein (Hirata et al., 2002; Bessho et al., 2003). They proposed a set of ordinary differential equations (ODE) describing their experimental observations:

$$\begin{aligned}
\frac{dx}{dt} &= By - Cx - Axz, \\
\frac{dy}{dt} &= -Dy + \frac{E}{(1+x^2)}, \\
\frac{dz}{dt} &= \frac{F}{(1+x^2)} - Gz - Axz.
\end{aligned}$$

where x and y are the concentrations of Hes1 protein and hes1 mRNA respectively, and z is the “Hes1-interacting factor“,that allows the system to have sustain oscillations. A-G are positive parameters that affect production and decay rates. x^2 represents a form of interaction of the Hes1 protein with itself.

Additionally, experimental evidence for cyclic expression within the PSM, falsified the 'cell cycle models' that attempted to explain segmentation clock phenomena. The cell cycle models of the clock assumed that the cell cycles along the anterior-posterior axis of the clock are in synchrony and that segmentation occurs when all the cells reach a certain time in their cycle (Primmett et al., 1989; Collier et al., 2000; McInerney et al., 2004). Baker and co-workers, revised their early cell cycle models and incorporated reaction-diffusion modeling that evolves along anterior-posterior axis in response to Fgf gradient signaling (Baker and Maini., 2007).

Following Hirata's approach, several ODE models have been put forth for the segmentation clock that considered cell autonomous oscillators and inter-cellular interactions, however, the ODE models did not include the real biological delays in transcription to mRNA and its translation to protein. In 2003, Nicholas Monk and Julian Lewis published very similar models for mouse and zebrafish respectively, proposing delayed differential equations (DDE) that included time delays and negative feedback loops for sustained oscillations with minimum number of parameters (Monk, 2003; Lewis, 2003). Further, Julian Lewis considered positive regulation by delta-notch activity combined with negative auto-regulation of Her1 and Her7 to generate sustained oscillations (Lewis, 2003). The simplest

DDE model can be written as:

$$\frac{dp(t)}{dt} = a\,m(t - T_p) - b\,p(t),$$
$$\frac{dm(t)}{dt} = \frac{k}{1 + (\frac{p(t-T_m)}{p_0})^2} - c\,m(t).$$

where p is the protein concentration in the cell produced with delay $\mathrm{T}_p > 0$, and m is the mRNA concentration produced with delay $\mathrm{T}_m > 0$. The positive parameters a, b and c are the protein synthesis rate per mRNA molecule, protein degradation rate and mRNA degradation rate respectively. p_0 is a critical concentration of the protein at which mRNA production is half its maximum value k. The period of oscillations is sensitive to the total delay; $\mathrm{T}_p + \mathrm{T}_m$.

Tiedemann and coworkers proposed a multicellular model encompassing an fgf8 gradient and Hes1 protein in the cytosolic and nuclear compartments, which enabled the system to exhibit sustained oscillation without requiring explicit time delays in the model (Tiedemann et al., 2007). While the idea of synchronization of cells within the PSM due to inter-cellular coupling was held (Jiang et al., 2000; Lewis, 2003; Horikawa et al., 2006; Holley, 2007; Ozbudak and Lewis, 2008), a popular mathematical model to describe synchronization, or more precisely, a large set of coupled phase oscillators (Kuramoto, 1984) soon found its place to explain segmentation clock dynamics. In this class of models, the PSM cell's 'state of expression' is represented by its phase. In 2007, Riedel-Kruse and coworkers, investigated such a model representation accompanied by inter-cellular coupling, using a mean-field approximation to study the effect of coupling strength on the synchronization properties of the system (Riedel-Kruse et al., 2007). Their study suggested that the coupled oscillators phenomena are an inherent property of the segmentation clock and a relatively high magnitude of coupling strength is required to overcome the intrinsic noise of the oscillators.

In another theoretical study, it has been shown that mutual entrainment of limit cycle oscillators at different frequencies coupled with a 'finite time delay' may provide several interesting synchronized solutions as opposed to the scenario without any delays, which has fewer solutions (Schuster and Wagner, 1989). Furthermore, Kuramoto models of coupled phase oscillatory systems with time delays have been widely explored to show the importance of delays in biological and chemical oscillatory systems beyond the segmentation clock. The measure of 'order parameter' in the Kuramoto models indicates phase synchrony among an ensemble of coupled oscillators (Strogatz, 1994; Yeung and Strogatz, 1999; Pikovsky et al., 2001).

In 2009, Morelli and co-workers, further extended the work of Riedel-Kruse et al., and formulated a delayed coupling theory to explain vertebrate segmentation. The theory captures coupling of phase oscillators, time delays in coupling and a frequency profile along the anterior-posterior axis considering the signaling gradients in the PSM (Morelli et al., 2009). The cells are spatially fixed with respect to each other and the clock is generated by phase oscillation frequency in the tail bud, while the wavefront mechanism moves posteriorly across the PSM, slows oscillations and eventually arrests the traveling phase wave in the form of somites. With a periodic sine function, phase differences among coupled oscillators are reduced. Patterns in the anterior PSM are realized at time intervals equal to the period of oscillations and the wave sweeps across a space, which represents the PSM tissue along the anterior-posterior body axis of the embryo to generate a striped pattern (somite). The phase dynamics of the coupled oscillators can be described by

$$\dot{\theta}_i(t) = \omega_i(t) + \frac{\epsilon_i(t)}{na^2} \sum_k \sin[\theta_k(t - \tau_i(t)) - \theta_i(t)] + \eta\zeta_i(t)$$

where the dot denotes time derivative, θ_i is the phase of oscillator i, ω_i is its intrinsic frequency, ϵ_i is the coupling strength, τ_i is the time delay in the coupling, η represents the noise strength and ζ_i is a random variable with zero average representing different noise sources. N discrete phase oscillators are arranged on a lattice, where each oscillator occupies a position $x_i = ia$ along the PSM axis and a is the characteristic distance between oscillators, i.e., the average cell diameter (Morelli et al., 2009).

Phase oscillator modeling offers great flexibility to explore the dependence on parameters and examination of analytical solutions. Moreover, it makes the mathematical analyses much simpler compared to DDE models, and allows easier examination of properties of the system, such as the effect of coupling strength and delays on the period of the clock or formation of somitic patterns. The theory suggests that the clock period is an effective collective period caused by signaling delays. The theory was experimentally tested and it was found that delta-notch mutant, *mib* had longer period with weakest somite boundary disruption phenotype, thus suggesting that coupling delay is strongly increased and the coupling strength is mildly affected in the mutant (Herrgen et al., 2010). Thus experimental evidence combined with theoretical formulation lead to the discovery of the first segmentation clock period mutant emphasizing that autonomous oscillator, coupling delays and coupling strength combinatorially regulate the collective period (Herrgen, 2008; Herrgen et al., 2010). The theoretical group has further shown that the segmentation clock is robust towards phase oscillators coupled with 'distributed delays' since the steady state depends on the mean of the delay distribution and is independent of its shape validating the discrete delay approximation originally assumed in the delayed coupling theory model (Wetzel et al., 2012).

The phase models elegantly take the biological complexity into account and yet have a simple formulation that allows experimental testing of clock parameters. However, the model descriptions are formulated for tissue level speculations that average out the cellular level details of the segmentation clock.

Recently, following the approaches of (Lewis, 2003; Hirata et al., 2002; Tiedemann et al., 2007) for modeling genetic oscillators within the PSM, Uriu and co-wrokers proposed to model using ordinary differential equations that takes into account the time evolution of: (i) *her* mRNA, (ii) Her proteins in cytoplasm and nucleus and finally (iii) the delta proteins expressed on the cell membranes. The model shows that traveling waves can be generated along the tissue with an appropriate choice of basal *her* mRNA transcription without incorporating fgf gradient signaling in contrast to some of the earlier studies (Uriu et al., 2009). The equations in details are:

$$\frac{dh_m}{dt} = \frac{K_1^n}{K_1^n + H_{pn}^n}(v_1 + v_c\hat{del_{pc}}) - \frac{v_2 h_m}{K_2 + h_m},$$

$$\frac{dH_{pc}}{dt} = v_3 her_m - \frac{v_4 H_{pc}}{K_4 + H_{pc}} - v_5 H_{pc},$$

$$\frac{dH_{pn}}{dt} = v_5 H_{pc} - \frac{v_6 H_{pn}}{K_6 + H_{pn}},$$

$$\frac{ddel_{pc}}{dt} = v_7 \frac{K_7^c}{K_7^c + H_{pm}^c} - \frac{v_8 del_{pc}}{K_8 + del_{pc}}.$$

h_m, H_{pn}, H_{pc} and del_{pc} denote the concentration of her mRNA, Her protein in nucleus, Her protein in cytoplasm and delta proteins expressed by neighboring cells respectively. For a more detailed description of the model, refer to (Uriu et al., 2009). Crucially, the cells were allowed to exchange their positions with their neighbors at random times in their model. The authors deduced from their work that exchange of positions of cellular oscillators with their neighbors considerably enhances their synchronization properties (Uriu et al., 2010). However, in this model, the cells were allowed to interact with neighbors through delta proteins immediately after changing their position. This assumption is probably not true for the PSM oscillators. Therefore, Uriu and co-workers re-modeled their theory in the form of coupled phase oscillators, in which they incorporated gradual recovery of intercellular coupling experienced by a cell after its movement. Through this work, they have provided a theoretical framework to study the effect of movement on synchronization properties of cellular oscillators within the PSM (Uriu et al., 2012).

In this work, we have experimentally measured three-dimensional movements rates of cells within the PSM and adapted the models proposed by Uriu and co-workers to derive properties of the system. These measurements would provide us for the first time scales of diffusion of cellular oscillators within the PSM. However, the task to obtain three-dimensional positions of nuclei over time requires optimizing image acquisition settings, image segmentation and tracking parameters. The section below gives a brief overview of the current challenges and techniques involved in quantitative live imaging.

2.7 Live imaging and image processing

During the development of multicellular organisms, cells undergo amazingly diverse changes with respect to gene expression, signal transduction, morphology, individual and collective motion (Karsent, 2008). Understanding how these processes are regulated and coordinated to build an embryo has benefited from recent advances in time-lapse fluorescence microscopy and the generation of a wide range of specific fluorescent probes. In combination, these have allowed spatial and temporal visualization of protein levels and localization, and hence cellular and sub-cellular dynamics, within developing embryos (Supatto et al., 2009; Megason, 2009; Long et al., 2009; Keller et al., 2008; Blanchard et al., 2009; Morrison et al., 2012).

The use of confocal microscopy allows high-resolution optical sections to be collected noninvasively at depths up to hundreds of microns within embryos. Visual inspection of these images allows a qualitative analysis and may be sufficient to assess dramatic effects in some systems. However, to fully characterize alterations in gene expression levels or cell motility, for example, quantitative evaluation is essential. Acquisition of embryonic images in three spatial dimensions over time (4D) with a balance between spatial and temporal resolution gives scope to convert image data into reliable quantitative measurements (Supatto et al., 2009; Roysam et al., 2006; Gerlich and Ellenberg., 2003; Eils and Athale., 2003). The computational image analysis required to accomplish this task is becoming a vital part of developmental biology, since it allows rigorous hypothesis testing and serves to bridge experimental data and mathematical modeling.

The primary task to detect an object in an image is done via segmentation. Following the object over time is achieved through tracking, where the key first step is accurate segmentation. Computer vision is well advanced in the field of image analysis, however, it is often either technically inaccessible or not easily modifiable for biological research purposes. Generic commercial image segmentation software packages are available, but these are expensive and do not allow access to their source code, meaning that they have to be used as black box. Furthermore, most biological imaging problems explicitly depend on the nature of the scientific question addressed, and therefore require tailored methods for image analysis to measure suitable quantities. As a result, development of more reliable and automated image analysis tools remains an important challenge. Several open source image segmentation techniques have been published, yet they may suffer in several aspects such as speed, reliability and ease of application. Importantly, objective measures of the accuracy of segmentation and straightforward protocols for optimizing the algorithms for specific imaging problems will foster wider use and increased confidence in the methods.

Finally, the task of quantitative imaging involves re-iteratively doing image acquisition to optimize spatial resolution in order to achieve reasonable image segmentation results that are further used for tracking objects. A number of parameters of acquisition directly determine the performance of tracking objects. For a good performance by an automated tracking algorithm, the acquisition time should be smaller than the displacement of the structure that is imaged (Meijering et al., 2012). Therefore conceptual feedback loops between image acquisition parameters, considering sample photo-toxicity and bleaching, good spatial resolution for image segmentation and high temporal resolution for tracking play vital role in accurate quantitative image analysis as seen in Fig. 2.10.

In this work, I provide an optimal framework for visualization of high quality cellular level image data with optimized spatial and temporal resolution for the PSM cells for the developing embryo. Subsequently, the data allows use of reliable quantification techniques that has been developed here. I have devised segmentation algorithm tailored for obtaining spatial positions of juxtaposed cells accurately, in a developing tissue followed by implementing tracking algorithms to obtain their positions over time. This further allows me to address specific questions about movements and oscillation dynamics at cellular resolution within the PSM of the embryo, and further enables interplay between relevant mathematical models and interpretation of experimental data.

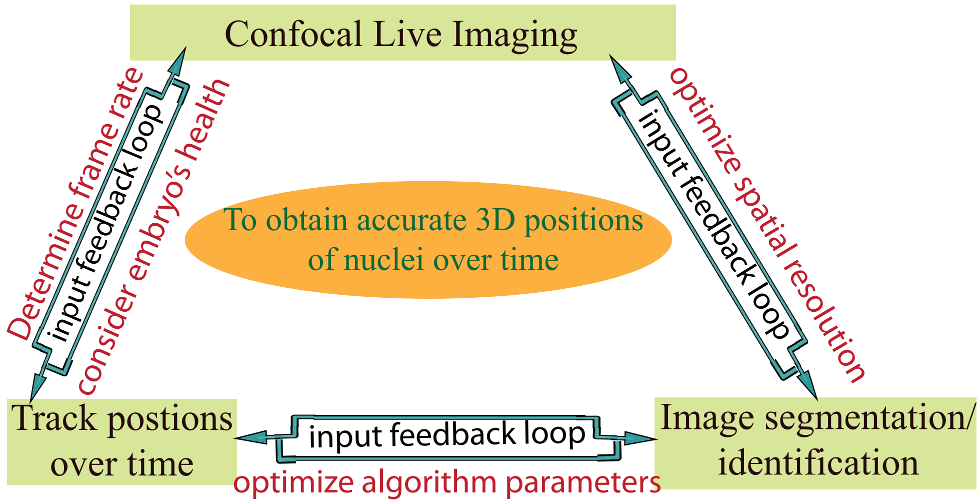


Figure 2.10: Interdependence between tools: Imaging, Segmentation and tracking requires optimization and trade-off between spatial resolution, time resolution and health of the live sample to obtain accurate image quantification.

Chapter 3

Aim of the thesis

Although several components of the clock have been identified, a complete understanding of the collective behavior of the genetic oscillators in the PSM is far from achieved. To explain and examine the relationship between temporal program and spatial pattern, we need to have a quantitative measurement of the properties of the system, in particular, the role of movements of cellular oscillators in the context of the clock. In this work, I seek to obtain quantitative data to understand how the movement of cells affects the spatio-temporal dynamics of cyclic gene expression that is regulated by inter-cellular signaling. I have the following aims:

- Establish imaging techniques to capture *in vivo* cellular dynamics of the moving oscillators within the PSM.
- Develop image analysis tools by formulating image segmentation and tracking algorithms to obtain precise tracks of single cells over time within the PSM.
- Use benchmarking techniques such as synthetic data and transplantation experiments to validate segmentation and tracking algorithm developed in this work.
- Quantify movement patterns within different regions of the PSM in WT and mutants, such as the effective diffusion co-efficient of cells.
- Subsequently, the tools become handy to extract phases of cellular oscilla-

tors of the PSM, which would provide scope to extract parameters, such as synchronization index and phase correlations for a population of oscillators.

In the following Chapter, I would outline my Results. In case any detailed explanation is required for better understanding of any section in the Results, I request the reader to kindly look into the corresponding section in the Methods Chapter.

Chapter 4

Results

4.1 Inspection of existing datasets for analyzing PSM cells

In order to monitor individual or group dynamics of cell movements and their oscillatory behavior, I first needed the spatial coordinates of cell positions within the PSM over time, i.e. precise tracks of individual cells. Around the time I started my thesis, Keller and co-workers published digital datasets of zebrafish embryo nuclear coordinates during 22.5 hours of development (for wild-type (WT) and *one-eyed pinhead* mutant, MZoep) (Gritsman et al., 1999) using a newly developed microscope, called a digital scanned light sheet microscope (Keller et al., 2008). The digital datasets of Keller et. al. were publicly available in the form of cell positions over time, however, not their tracks. I analyzed the dataset and concluded that the data quality was insufficient for our purposes. We were particularly interested in the PSM cells at later stages of embryo development (about 16 to 20 somite stage), where the number of cells is large, and the cells are closely juxtaposed and are mixing at varying rates during rapid and dynamic morphological changes in the tissue shape over developmental time. It seemed a technically challenging task to digitally recognize and extract just the PSM cells alone, from a dataset for all embryonic cells during development. However, the analysis provided an excellent opportunity to understand and address the goals and challenges involved in optimizing experimental details affecting image acquisition, segmentation and cell tracking variables. Thereafter, I went on to set up an experimental framework for obtaining high quality spatio-temporal resolution data by making the appropriate choice of microscope and exploring sample preparation methods to image live zebrafish embryos, with minimized bleaching and photo-toxicity issues.

4.2 Exploration and establishment of image acquisition techniques

I first sought to explore a range of microscopes supported by the Light Microscopy Facility (LMF) in our institute, choice of transgene, sample preparation and mounting techniques.

Choice of transgene: Our goal was to determine precise cell positions over time within the PSM and since the nucleus is approximately the centre of the cell, it seemed a viable option to choose a nuclear marker to read cellular positions. Therefore, I used a chromatin histone protein fused with green fluorescent protein (gap) or a similar m-cherry variant (Hans-Hermann and Kaether, 1996; Chudakov et al., 2010). The gfp fusion zebrafish transgenes seemed our potential candidate to visualize fluorescent labeled nuclei and nuclear divisions, as readouts for cellular positions over time.

Sample preparation: The sample preparation method used in this work is a variant of the technique that has been previously described in (Herrgen et al., 2009) and has also been used to precisely characterize zebrafish somitogenesis period (Schröter et al., 2008). I used a similar sample preparation method to acquire high quality images at cellular resolution. A petri-dish coated with agarose with conical depressions such that the de-chorinated embryo head can fit laterally into the depression while the posterior-body develops parallel to the agarose, has been used in this work for imaging as described in the methods section 6.3 and illustrated in Fig. 6.1A. The embryo develops under the microscope placed in a petri-dish, filled with E3 medium (that is also used for raising embryos otherwise) and grows in in its natural environment without any external pressure.

Choice of microscope: The confocal microscopy technique seemed to be the best option since it has the capability to provide high cellular resolution three-dimensional images. The width of the confocal pinhole determines the amount of light that reaches the sample and removes any out of focus light from the image.

As a result, high quality, sharp two dimensional images can be obtained. A series of two dimensional images acquired over a certain depth (along the z-direction) can be put together to give a good three dimensional image (Wright and Wright., 2002; Amos and White., 2003). I took care to optimize the number of imaging hours and problems of photo-toxicity that would directly affect health of imaged embryos and might cause slower development and in turn cause abnormal movement of cells. For this reason, the images were not acquired for more than 2 hours. Typically, images of 512×512 in x-y and 20 - 30 slices in z respectively, with voxel size of $0.691 \times 0.691 \times 1.75 \mu m^3$ were obtained. Furthermore, the imaged embryos were carefully removed from the agarose and their development was verified until 30 hours under normal physiological conditions.

Selective Plane Illumination Microscope (SPIM), a technique that has recently gained popularity, was also explored in this work due to its capability to acquire images at high speed with fairly good resolution, lower photo-toxicity issues and greater flexibility for mounting samples (Huisken et al., 2004; Huisken and Stainier., 2009). However, the first hand images were obtained on a test SPIM microscope provided by Zeiss with several problems regarding mounting live embryo and image blurring. Hence, at that time, the technique was not explored any further, and the image processing pipeline established here was designed for images obtained from a Carl Zeiss upright confocal microscope, with a 40x, 1 NA dipping lens. We later revisited the SPIM together with Dr. Jan Huisken and demonstrated that we can achieve convincing results with the image processing algorithms developed through this work. The PSM is a bilateral tissue, therefore, we chose to image either lateral sides of the developing embryo. The choice of an upright microscope was compatible with our sample preparation technique.

4.3 Image processing pipeline

The general workflow for quantitative image analysis implemented in this work is shown in Fig. 4.1. As mentioned earlier in section 2.7, the task of quantitative imaging involves re-iterative combination of steps from image acquisition to tracking objects. The idea was to establish an imaging set up that would

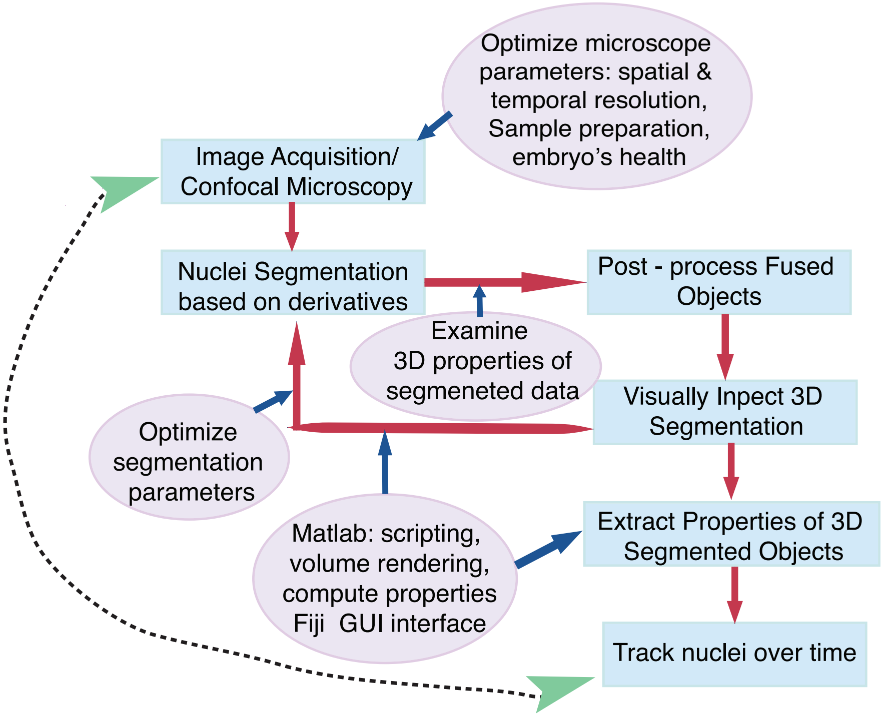


Figure 4.1: Blue boxes in the flowchart give a step-by-step procedure used in this work. Purple ellipses indicate the important guideline and parameters required for optimizing the image processing pipeline at different stages.

allow us to track individual nuclei in 3D without compromising the health of the embryo. Here, I have obtained high resolution three-dimensional stacks of the PSM over time using an upright confocal microscope followed by applying a three-dimensional segmentation algorithm based of image derivatives. The three-dimensional segmentation algorithm developed in this work is a fast and accurate technique that is able to segment about $\sim$ 2000 nuclei within the PSM tissue, that undergoes rigorous morphological changes and has juxtaposed cells that exhibit continuous diverse and dynamic cell motions during development. Thus, our PSM image data provides a technically challenging platform for image analysis. To tackle this, I designed a segmentation algorithm comprising of a set of sequential steps, that ultimately is able to segment nuclei with high accuracy in 3D. Further on, I show the use of artificial data and for the first time the use of transplant experiments to validate the 3D segmentation algorithm *in vivo*. Due to high accuracy of segmentation, I have been able to use a tracking algorithm based on nearest neighbor centroid detection to follow nuclei over time. The results are further used to characterize the movement rates in different regions of the PSM for WT and mutant backgrounds. It was also possible to observe single cell *in vivo* protein oscillations using the method developed here.

4.4 Imaging the growing pre-somitic mesoderm in 4D

I acquired three-dimensional image stacks of the PSM using an upright confocal microscope, illustrated in Fig. 6.1A. The 40x dipping lens is immersed into a petri-dish such that the objective is focussed on the lateral side of the developing PSM, after the last formed somite as shown in Fig. 4.2A (Supplementary movie: S1 with movie legend in Chapter 7). The zebrafish embryo PSM stack is obtained over several parallel xy-slices of the tissue, from nuclei of the outer epithelial skin layer through the PSM, until the notochord nuclei are visible. Fig. 4.2B. shows an image of a typical three-dimensional stack (one time point) obtained in such a manner and 4.2C. shows snippets of three two-dimensional slices from different confocal planes of the imaged tissue (Supplementary movies: S2-A, S2-B, S2-C, with movie legends in Chapter 7). A range of image resolution was tested in the z-direction in order to achieve good 3D segmentation of nuclei. Hence, the next goal was to design an appropriate algorithm to acquire individual 3D positions of nuclei accurately over time.

4.5 3D nuclei segmentation within the PSM

Advances in imaging techniques, fluorescent markers and innovative sample mounting ideas are providing rich information regarding biological systems across several scales. With the advent of such technologies to collect image data, it simultaneously becomes important to have reliable quantification techniques along with visual inspection. The problem of nuclei segmentation in a crowded, developing embryonic tissue is not at all new, however the topic is still under great scrutiny, since we have not yet been able to achieve satisfactory results in terms of accuracy and reliability in image segmentation.

The great diversity of biological questions and the physical scales involved makes the choice of segmentation method highly problem dependent. A major drawback

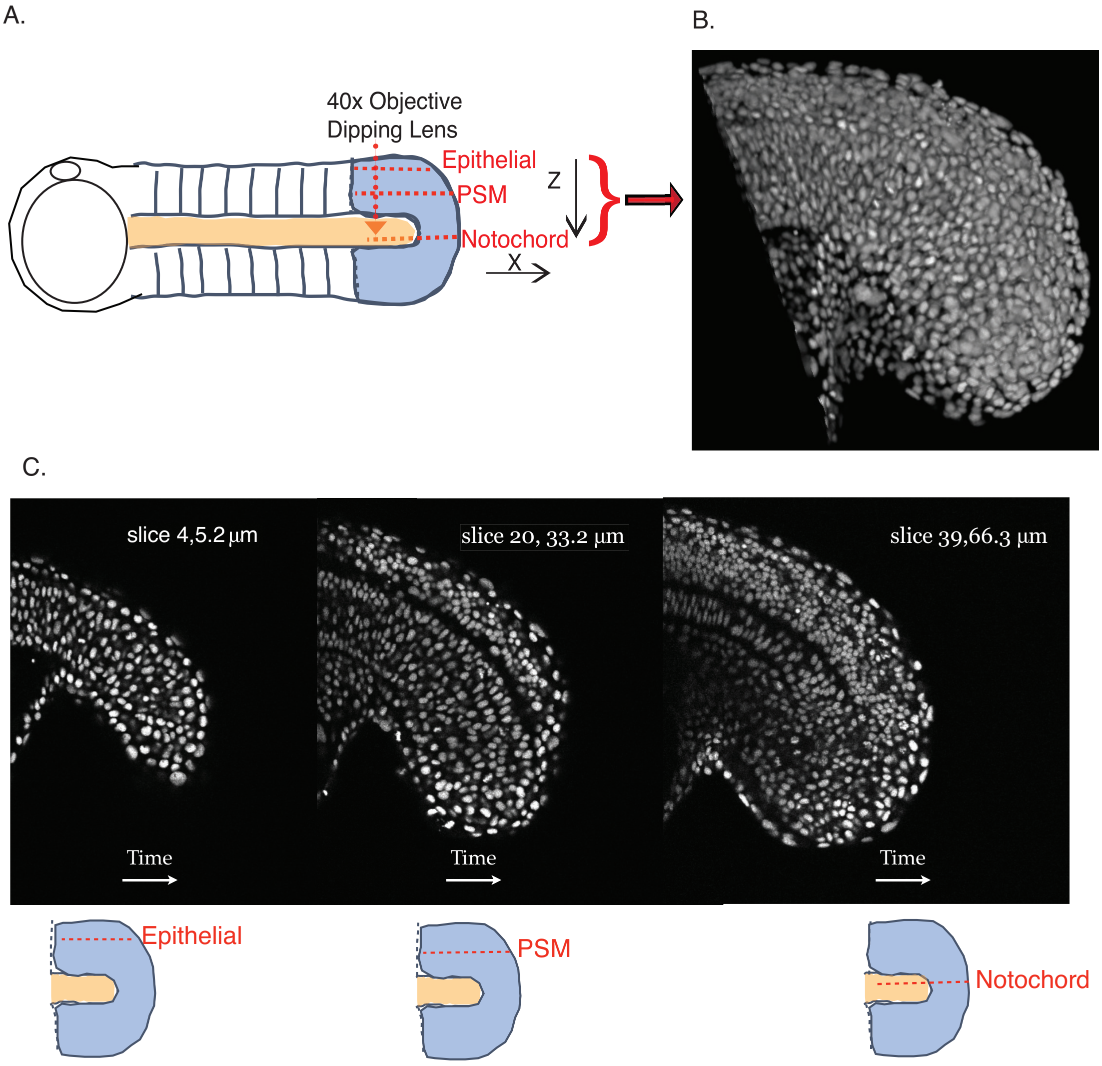


Figure 4.2: A. Cartoon view of the zebrafish embryo. 40x objective is focussed on the lateral PSM tissue with tissue slices encompassing epithelial nuclei, PSM nuclei and notochord nuclei of an 18 somite stage embryo. B. A typical three-dimensional view of the stack obtained from Fiji, 3D viewer. C. Snap shots of confocal slices of the PSM at different depths indicated in microns.

in the field of image segmentation is the lack of measure for performance that could rate all the algorithms under a single umbrella, and thereby would allow users to trust the effectiveness of algorithms over variety of available biological image datasets. A number of segmentation algorithms have been proposed, many of them focussing on segmentation of migrating cells *in vivo* as well as *in vitro*. Image analysis tools based on active contours compute image segmentation by evolving a contour in the direction of the negative gradient of the image energy, however, such methods are sensitive to fuzzy boundaries and initial conditions (Kass et al., 1988; Osher and Sethian, 1988; Meegama and Rajapakse, 2003). Methods using region based active contours overcome such flaws because they do not require explicit parameterization or topology constraints, like snake contours (Chan and Vese, 2001), however they are computationally expensive. The active contour methods are quiet sophisticated and adaptive compared to the standard thresholding techniques. Thresholding methods include a binary operation that is a fast way to achieve image segmentation. A threshold value, say, T is set for a particular image. All intensity values above T are assigned to be part of foreground of objects (of value 1) and below the value are supposedly part of background intensities (of value 0) (Brakenhoff et al., 1995; Sezgin and Sanku., 2004). However, to get accurate results using simple thresholding can be challenging because thresholding cannot separate touching objects properly. To overcome this problem, watershed algorithm and several of its variants have been proposed, that can clearly separate touching objects, but over-segmentation is likely to happen simultaneously (Meyer, 1994; Bleau and Leon, 2000; Lin et al., 2003; Gonzalez and Woods, 1992). The gradient flow tracking method can overcome this problem, however it may find difficulty in processing textured cells (Li et al., 2007). My challenge was to develop an algorithm that has the ability to separate touching nuclei correctly in 3D space and that can be easily used to process several embryos varying in nuclei number, nuclear size, volume and density. Here, I propose a segmentation algorithm that combines a sequence of steps, based on well known mathematical functions, until high accuracy of segmentation is achieved. The algorithm is developed specifically to address the problem of touching nuclei images in 3D in developing tissues. Even though we have acquired images with high spatial resolution (while compromising for temporal resolution), the major task still remains to separate closely spaced nuclei. The problem grows severely with developmental time in embryonic tissues as the number of nuclei is increasing, and they emerge in varying sizes and shapes, their morphology and mobility being tightly attributed to the tissues and organs they are destined to form.

Typically, grayscale images of 8-bit (range of values between 0 and 255) or 16-bit

(range of values between 0 and 65535) are acquired wherein high values represent intra-nuclei regions or foreground, while medium ranged values could be a part of inter-nuclear regions with Poisson noise in the proximity of nuclei, and even lower values represent background intensities. The borders between the three qualitative categories mentioned here entirely depends on the image acquisition parameters, microscope and the qualitative spectral properties of fluorescent tagged proteins. I propose a combination of image derivatives to crown every nucleus in the image that can be easily converted to a binary image using the Otsu's thresholding method that yields several properties of the 3D segmented objects, such as, object volumes, voxel lists and centroid of voxel lists. These properties are further used to enhance segmentation performance and hence boost the overall accuracy of the method proposed in this work.
When we look at a two-dimensional slice from a 3D stack as seen in Fig. 4.3A., and apply a simple Otsu's thresholding method (Otsu., 1979), we see that the contour of thresholding does a fairly good job of separating nuclear images. However, in regions where the spatial inter-nuclear distance is small, the thresholding method fails substantially as seen in Fig. 4.3C. Although, we could achieve good xy resolution using a confocal microscope, simple thresholding still leads to errors in 2D. Furthermore, errors grow in 3D, owing to the axial resolution of the confocal microscope, which is poorer than the xy resolution.

In the following, I describe the algorithm, as outlined in Fig. 4.4 that addresses this problem. We start from three-dimensional z-stacks of 2D confocal x-y image slices that are finally segmented in 3D. In essence, the algorithm first smoothens the raw input images to eliminate noise within the images, followed by applying a series of masks based on image derivative steps that clarify nuclear edges in 2D. The resulting 2D image is segmented by binary thresholding, and assembled into a 3D image by connecting like pixels in neighboring slices. The same algorithm can also be formulated for 3D image volumes. However, it is easier to pictorially demonstrate the intention of each step in 2D. Therefore, we focus here on the version for 2D slices, that are later connected to give 3D segmented objects. The algorithm has been titled 'CWNT' meaning crown wearing nuclei tracker.

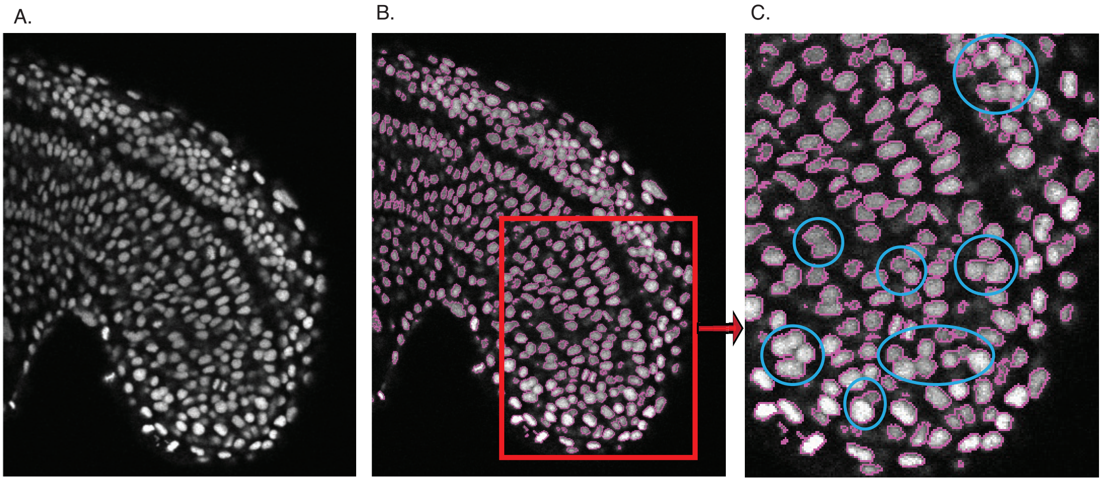


Figure 4.3: A. Lateral view of a raw, grayscale PSM image slice (x-y) from a three dimensional stack of 18 somite stage embryo. B. Contour plot of Otsu's thresholding plotted over original raw input image C. Image regions highlighted with fused nuclei in blue circles where Otsu's thresholding fails.

4.5.1 Image De-noising

Intensity fluctuations, both in background and in intra-nuclear regions severely affect image segmentation results. To reduce this perturbation before edge detection, we first apply linear and non-linear de-noising filters, as given by equation (6.1). Typically, we used a low pass Gaussian filter, which removes high frequency noise. If the acquired image is severely affected by noise, we employ de-convolution and the median filters following Gaussian filter. The choice of noise removal filter strongly depends on the quality of the image obtained. We have listed a few linear and non-linear filters in section 6.1, that might be appropriate to use depending on the noise of the image. To further remove intensity fluctuation in the image, whilst preserving edges of nuclei, I applied the non-linear isotropic diffusion filter proposed by Perona and Malik (Perona and Malik., 1990). It smoothens image regions less that have high gradients and therefore potential edges compared to low gradient regions. The details of implementation is illustrated in Fig. 4.4, description is as per equation (6.2). The diffusion coefficient implemented in this work (described in section 6.4.2) exponentially decreases as the gradient of signal intensity becomes steeper.

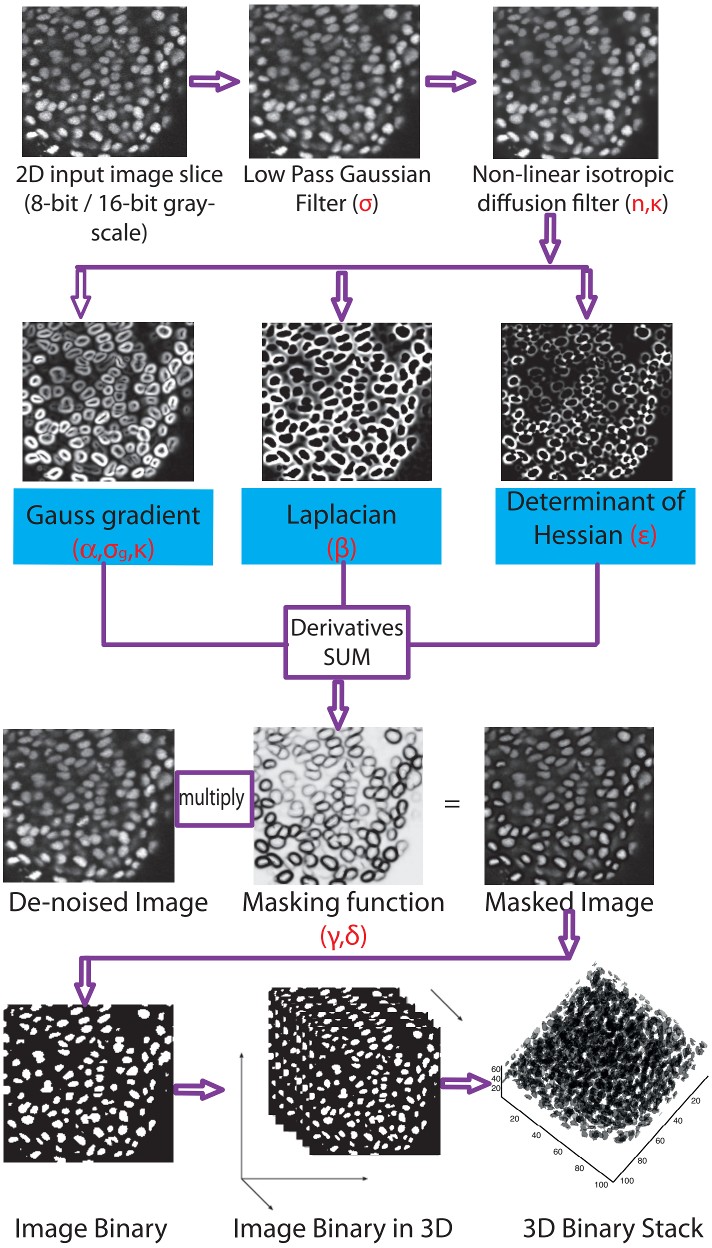


Figure 4.4: Segmentation algorithm, CWNT, described methodologically for a 2D raw image slice (grayscale) from a stack. Gaussian blur is applied to a 2D slice, followed by a non-linear isotropic diffusion filter that smoothens pixels inside of the nucleus, whilst preserving edges. First derivative of the image using a Gauss gradient emphasizes the edges, Laplacian marks the maximum of the gradient and Hessian highlights the saddle points and cuts touching spaces between nuclei. Image derivatives are combined with tunable parameters (indicated by red greek glyphs) into a tangent hyperbolic masking function that allows tuning the steepness of edges. This masking function is applied on the de-noised image, such that nuclear edges are masked. Otsu's thresholding is applied to the masked image. Similarly, all slices of the stack are converted to binary image. 2D binary slices are connected based on similarity of like pixels to give a 3D binary stack (surface rendered gray objects).

4.5.2 The core segmentation algorithm

Following the de-noising steps, the task of the core segmentation algorithm is to find the edges of the nuclei, in particular those edges where nuclei are touching. I computed first and second derivatives of the signal intensity in each 2D slice to detect edges of nuclei. Signal intensity steeply increases from the edges of nuclei toward the center, while it is almost flat around the center. Therefore, I computed the magnitude of the gradient and the Laplacian of the signal intensity to enhance nuclei edges as given by equations, (6.4) and (6.7). The magnitude of the gradient increases at the boundaries between a nucleus and the background and it takes a maximum value at the point where the slope of the signal intensity is steepest i.e. the inflection point. At the inflection points, the Laplacian of the signal intensity changes its sign. The key innovation in this image processing algorithm is to be able to infer nuclear edges where neighboring nuclei are touching each other. To achieve this, I applied at each pixel, the negative determinant of the Hessian of the image, which marks the saddle points of the image, where nuclei closely touch each other, for details see section (6.8).

I finally combined the derivatives in the form of a masking function that highlights nuclear edges. For this, I first added all three forms of the derivatives: magnitude of gradient (ADF), Laplacian (L) and the determinant of the Hessian (H), with different weighing constants; α, β and ϵ as given by equation (6.9). I then combined the sum into a hyperbolic tangent function that detects the edges of the nuclei. The masking function is zero at edges and touching points between nuclei, and otherwise one at all points. The product of the masking function and the de-noised image produces an image with suppressed signal intensity at touching points of nuclei, ideal for thresholding, given by equations (6.10), (6.11) and (6.12).

4.5.3 Thresholding and generation of 3D images

I obtained binary images by applying Otsu's thresholding method to the masked image. The segmentation protocol was similarly applied to all 2D slices and similar pixels from the neighboring slices were combined to give 3D binary objects as seen in Fig. 4.4.

The power of our nuclei segmentation method is illustrated in Fig. 4.5. Contour plot of Otsu's thresholding (Otsu., 1979) applied on the raw input image is unable to separate closely spaced nuclei, as seen in Fig. 4.5A., whereas Fig. 4.5B., clearly illustrates the effectiveness of using our image segmentation method, which easily allows us to separate closely spaced nuclei. The resulting threshold contour matches well to the apparent edge of the nucleus in the raw image. The computation time of image derivatives takes less than a minute for a stack of about 40 slices (70 µm thick) with Matlab R2011a on a 4-core machine.

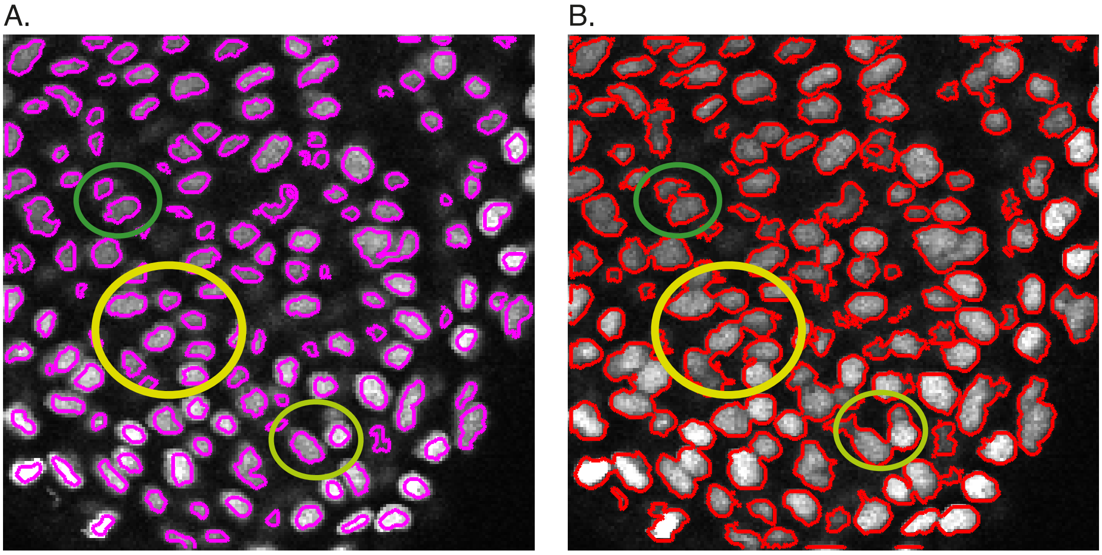


Figure 4.5: Contour Plot of Otsu's thresholding in 2D for all nuclei demonstrates the effectiveness of the segmentation algorithm. A. Threshold contour (pink) for processed image using the segmentation algorithm. B. Threshold contour (red) for original 2D slice,unable to separate fused nuclei. Similarly colored ellipses highlight the same regions in both the plots.

The segmentation steps demonstrated so far, attain reasonably good results, as seen in the 3D binary image obtained after segmentation in Fig. 4.6. Supplementary movie: S3-A shows 3D movie of the embryo over time as an input to the segmentation algorithm and Supplementary movie: S3-B shows the same movie after image segmentation (movie legend in Chapter 7). Each segmented object is now defined by three properties: voxel list (voxels in x, y, z comprising the object), volume (total number of voxels) and centroid of voxels. The position of each nucleus in the image volume is given by the centroid of the object. The

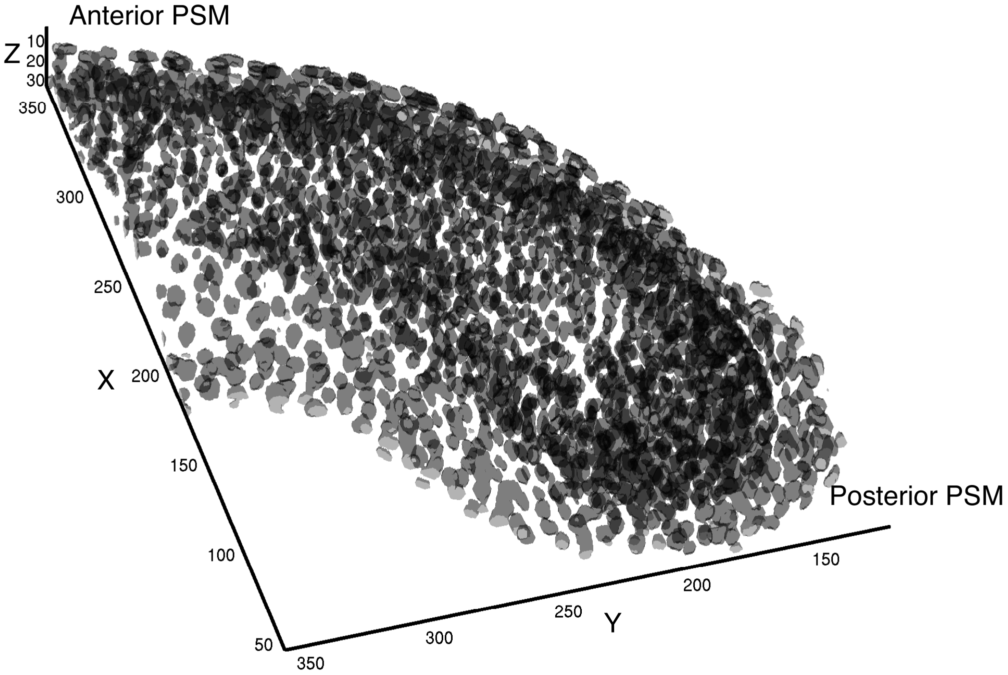


Figure 4.6: Surface rendered 3D binary image for a 19 somite stage embryo after successful segmentation. Rendered nuclei indicated in grey for a stack of the PSM

voxel list of the 3D images is used to calculate spatial properties of the segmented objects, and objects that exceed empirically established volume limits are subject to post-processing to separate inappropriately fused (under segmented) objects. For simplicity, 3D segmented data clip shown in Fig. 4.4 has been used here to demonstrate post-processing steps.

4.5.4 Post-processing resolves under-segmentation

I examined the segmentation performance by analyzing the properties of the segmented objects. I used the histogram of the object volumes as a tool to draw a boundary between correctly segmented and under-segmented objects with an empirically determined cut-off for fused objects (for details refer to section 6.4.6).

The histogram plot of the 3D volumes of segmented objects depicts the range of nuclear volumes within the data including noise (voxels less than 5), correctly segmented volumes and potential fused candidates, as indicated in Fig. 4.7. While the image segmentation steps listed above have solved most of the problem depicted in Fig. 4.5, some residual fused objects still remain, particularly in the z-direction that indicate quite large volumes of under-segmented objects, comprising multiple fused nuclei. To circumvent this problem, one option for splitting fused objects was to vary the values of parameters contributing to the masking function in equation (6.11). Increasing the weights of the parameters would results in erosion of the nuclei at the edges that consequently would reduce under-segmentation and hence improve 2D nuclei identification. However, this approach has the disadvantage that the volume of the nuclei is strongly compressed. To simultaneously reduce under-segmentation and preserve as much of the native morphology as possible, it proved useful to apply a post-processing step to the 3D segmented images, in order to correct segmentation mistakes in the form of fused nuclei. It is further noted from the histogram plot that the image comprises of a range of 3D nuclear volumes that are correctly segmented, therefore, mean volume cannot be used to determine the number of fused nuclei directly.

I put forward two post-processing methods based on unsupervised clustering, namely k-means and Gaussian mixture models (GMM) that locally correct for segmentation mistakes (Hartigan and Wong, 1979; Figueiredo and Jain, 2002; McLachlan and Peel, 2000; Press et al., 2007). K-means groups input data points into clusters, that largely depends on distances between the points (3D voxel list coordinates in this case). Whereas, GMM fits mixture of several Gaussians with different means and variances to data points, and finds best suited parametric models, when combined with Akaike information criterion (AIC) and thereby determines clusters in data. The post-processing methods operate locally on each potentially fused object and find new reasonable clusters of voxels and their respective centroids, thereby enhancing segmentation performance considerably. A typical correctly segmented object/nucleus volume has a unimodal frequency distribution of voxels in each x, y and z direction (from the voxel list), and has a single centroid for a given volume as seen in Fig. 4.8. Thus, our post-processing

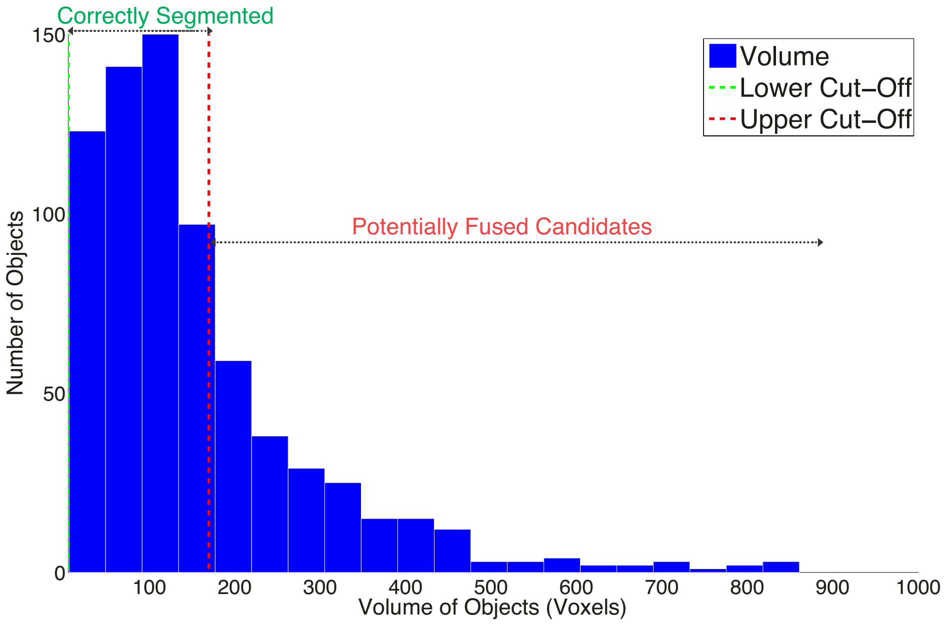


Figure 4.7: Histogram plot of segmented object volumes within a single stack (one time point) allows preliminary segmentation assessment. A lower cut-off of 5 (indicated by the green dotted line) excludes object volumes less than 5 voxels considered noise. The upper cut-off (indicated by the red dotted line) is determined empirically (see 6.4.6) based on the mean and standard deviation of the segmented volume list. Objects with volumes above the upper cut-off form potentially fused candidates that are further processed for segmentation corrections.

steps applied to correctly segmented objects do not alter the centroid position, indicating that our local re-segmentation approach is not biased towards splitting objects unlike the classical watershed that often leads to over-segmentation.

However, the corresponding voxel distribution in x, y, z for each outlier volume (or fused objects) indicate the number of fused nuclei. These objects are separately analyzed by finding local maxima or 'total number of peaks' in each voxel list and this information is fed to the k-means clustering algorithm to re-assign voxels. The distance between two consecutive peaks approximately depends on the voxel sizes of acquired image. We typically defined voxel sizes of $0.691 \times 0.691 \times 1.75 \mu m^3$ and for this resolution, it was estimated that a nucleus usually spans about 3µm in x and y and approximately about 2µm in the z direction. We considered the span of a nucleus to be approximately equal to the minimum

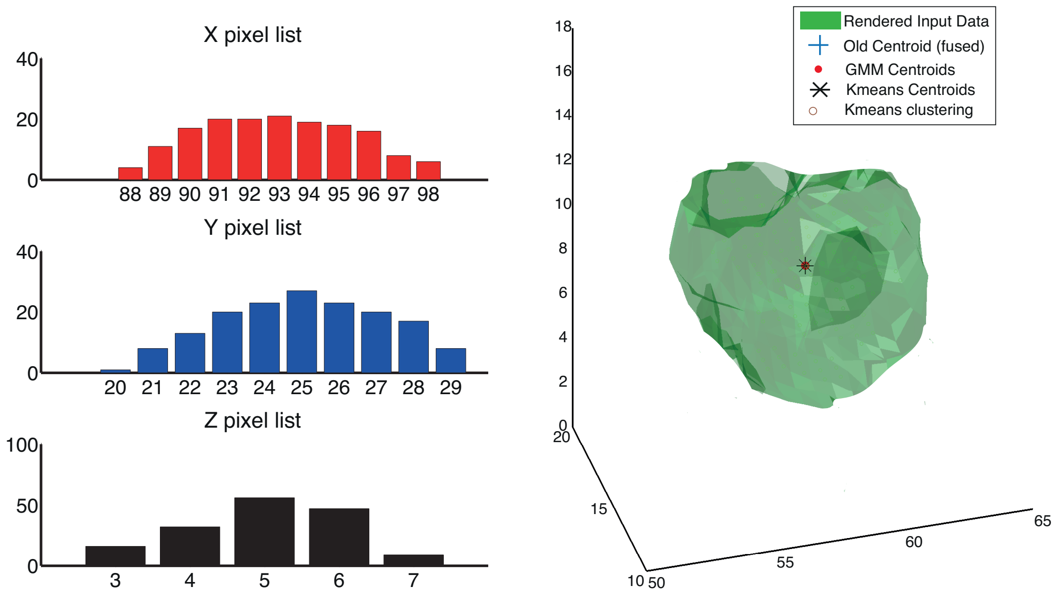


Figure 4.8: A correctly segmented object of 156 voxels with unimodal distributions in x,y and z directions. The rendered input data for a single nucleus with centroid plotted on top. Centroids generated after segmentation and post-processing steps (GMM and k-means). Post-processing does not alter the position of the segmented centroid for a correctly segmented object.

peak distance in voxel distribution plot for each direction. Therefore, this further helped in setting the minimum distance criteria between two consecutive peaks in the voxel distribution, which is a read out for number of clusters fused together. Minimum peak distance was set to be about 4 for x and y voxel distributions and about 1 for the z direction. The result of k-means heavily depends on the number of local maximas found in the frequency list. On the other hand, GMM is initialized using the local maximum peak from the voxel lists. GMM method fits 3D Gaussians clusters to the voxel lists. However, GMM combined with AIC minimizes the number of Gaussian clusters based on the features of the distributions.

Both post-processing steps, k-means and GMM insightfully utilize the features of segmented objects to split fused nuclei and thereby correct for segmentation

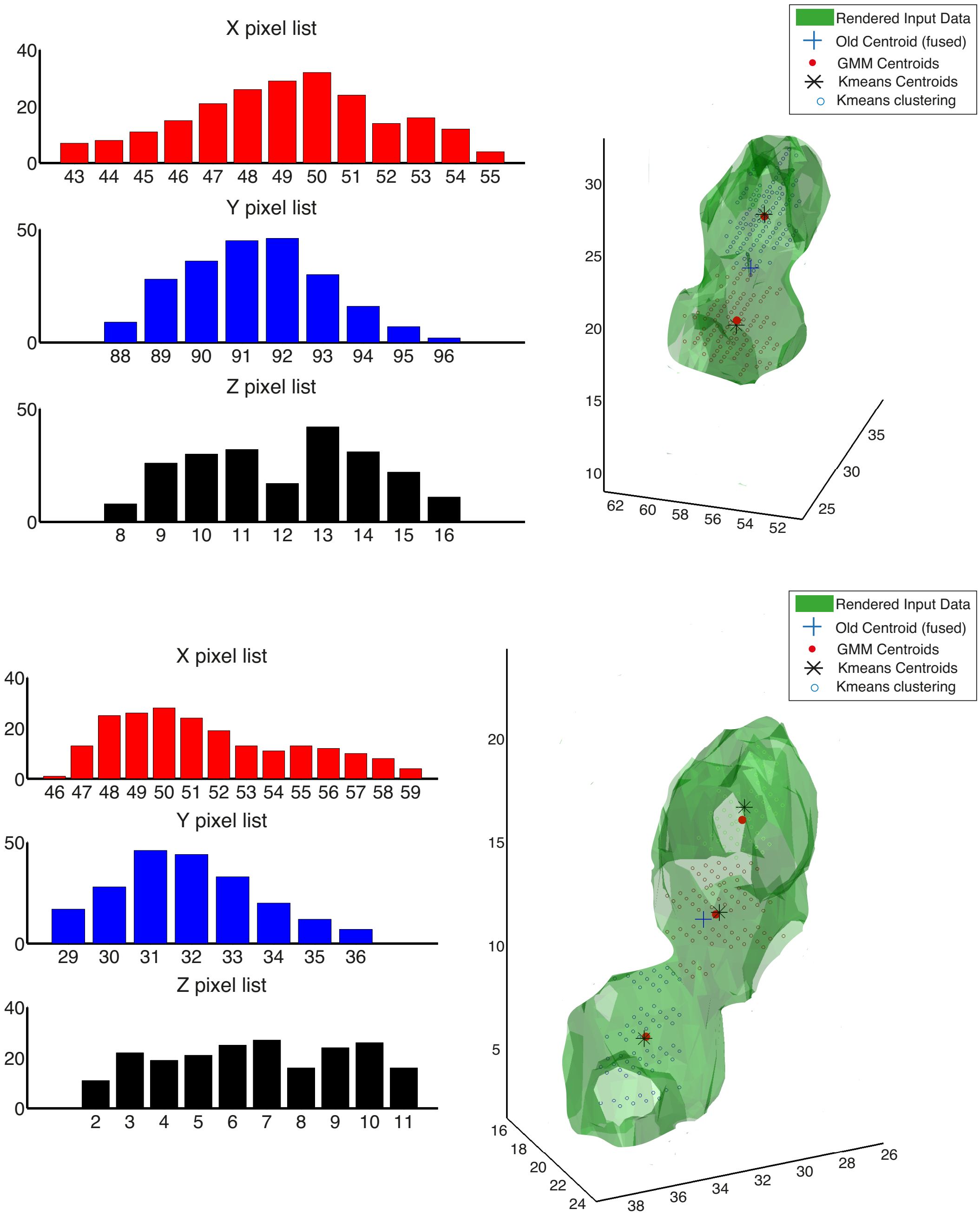


Figure 4.9: Fused nuclei examples. Bar plots of sampled voxel list exhibits unimodal distributions of peaks in x and y directions, a bi-modal distribution in the z voxel list (top panel) and a tri-modal distribution (bottom panel) comprising the object. Post-processing steps are applied to correct segmentation mistakes. Here, both GMM and k-means give similar results, by splitting objects into new volumes and their centroids. In top panel, GMM splits into 108 and 111 voxels while k-means gives 113 and 106 voxels respectively, for the two new objects. In the bottom panel, both methods split a volume of 207 voxels into three new objects with 66, 75 and 66 voxels respectively.

mistakes. The frequency distribution plot for the voxel list in each direction for fused objects would exhibit multi-modal distribution, indicating the number of clusters of nuclei fused together. This information is used as a first step to feed into the clustering methods (for details, refer 6.4.6) to initialize splitting nuclei. Fig. 4.9 shows multi-modal frequency distribution within the z voxel list and a roughly unimodal distribution in the other two directions. Both post-processing steps arrive at similar results in case of two or three nuclei fused together and further show considerable overlap in the position of centroids found, as illustrated in Fig. 4.9.

Local noise within the nucleus is also a direct observable in the sampled voxel list that could feign local maximas that are potential indicators of number of clusters, and therefore it is vital to pertinently de-noise the images. The analysis in this work indicates that k-means method is more sensitive to local noise in the form of local peaks that appear in the voxel list frequency distribution, and hence may result in erroneous number of nuclei. However, unlike k-means, GMM is only initialized with the number of clusters from the voxel list, and when combined with AIC, GMM elegantly fits correct number of nuclei models to the list, and therefore is observed to be less sensitive to small local peaks (noise) in the voxel lists. However, both approaches aim for separating fused nuclei and gave comparable results.

The stem plots after segmentation and post-processing steps indicate the effectiveness of applying proposed methods to rectify segmentation errors locally as illustrated in Fig. 4.10. Number of fused nuclei has considerably reduced after implementing post-processing steps compared to what was achieved directly after segmentation as seen in Fig. 4.7.

Image segmentation is a challenging problem and several algorithms have been proposed in the literature, however, it is yet harder to obtain high accuracy. Nevertheless, improvements on segmentation is extremely valuable since it has direct impact on tracking performance. Here, we have proposed an approach that can automatically segment nuclei with varying volumes, obtained from a real biological tissue. Additionally, the algorithm also allows the options of verification check points in the form of histogram plot of volumes and 3D rendering of objects with their frequency distribution plots of voxel lists, making this approach of image

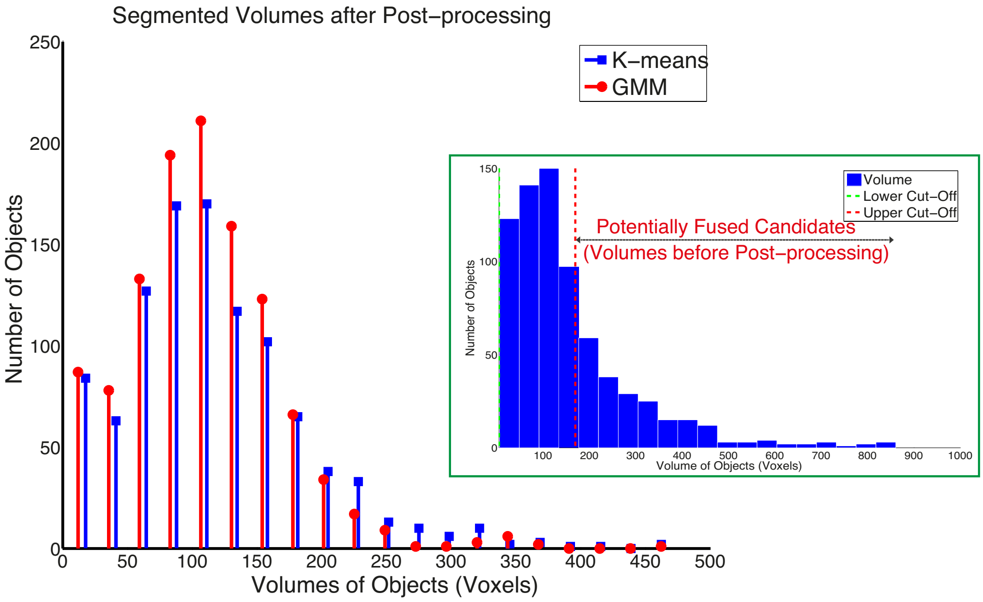


Figure 4.10: Stem plot of volumes of segmented objects after post-processing fused nuclei. GMM method achieves superior performance in splitting objects in comparison with k-means. Right box shows fused candidates segmented volumes before post-processing same as seen in Fig. 4.7.

segmentation reliable in the context of live imaging of embryonic tissues. The algorithms were developed in Matlab, and are also available in the Fiji library (Schindelin et al., 2012), as a user-interactive plugin, written by Dr. Jean -Yves Tinevez.

4.6 Verification of nuclei segmentation algorithm

Segmentation algorithms for biological research are specifically built depending upon the markers used and questions asked for understanding biological processes. The problem of image segmentation can be solved using several techniques. Edge detection algorithms based on image derivatives have high speed, however, they may be sensitive to false positives and therefore mis-classify edges. Here, I have provided a framework, which goes through several steps, first at coarse grain level and later, finely tunes the image parameters locally to achieve high quality results. To do this task, I have put forth the use of refined information from one step into the next, until error rates in image segmentation decrease considerably. However, it is still difficult to asses the performance of algorithm with respect to our image data. The missing key here is the verification steps, that allow easy assessment or ranking of algorithms. Here, I propose the use of verification steps in order to rate the algorithm performance.

4.6.1 Assessment of segmentation algorithm with Imaris

Although, we precisely know the number of nuclei during early stages of embryo development (Warga and Kimmel, 1990), the exact number of nuclei within the PSM and its change at late developmental time is much less certain. Here, I first used Imaris (Bitplane), an imaging processing software to qualitatively assess the CWNT algorithm performance. The 'Segmentation-spots' feature from Imaris was used for this purpose and all the nuclei were first segmented in a semi-automated manner. Fig. 4.11A. shows the 3D rendered nuclei of the PSM for a single time point, with the centroids found by the Imaris software. I chose suitable parameters based on visual inspection (using volume rendering feature) to obtain accurate segmentation for all the nuclei in 3D for the first time step. The same set of parameters was used for all later time points. Later, segmentation mistakes had to be corrected manually on the software. Next, I used the CWNT algorithm to segment nuclei for the same stack and all time samples. Fig. 4.11B. provides a qualitative measure of performance of the two algorithms. The total number of nuclei given by the two algorithms is fairly similar, as indicated in the plot. It gradually decreases in both the cases owing to the fact that the embryo rolls out of the imaging frame during its development. This is illustrated in Fig. 4.11B by outlining the 3D snapshot of the PSM (3D viewer in Fiji) at the first and last time points. Supplementary movie S4 clearly highlights this feature, movie legend in Chapter 7. We next evaluated the performance of the two algorithms using artificial datasets.

4.6.2 Segmentation Efficacy with artificial data sets

Synthetic images were generated in order to systematically test the CWNT algorithm for different signal-to-noise ratios (SNRs) in the image and density of

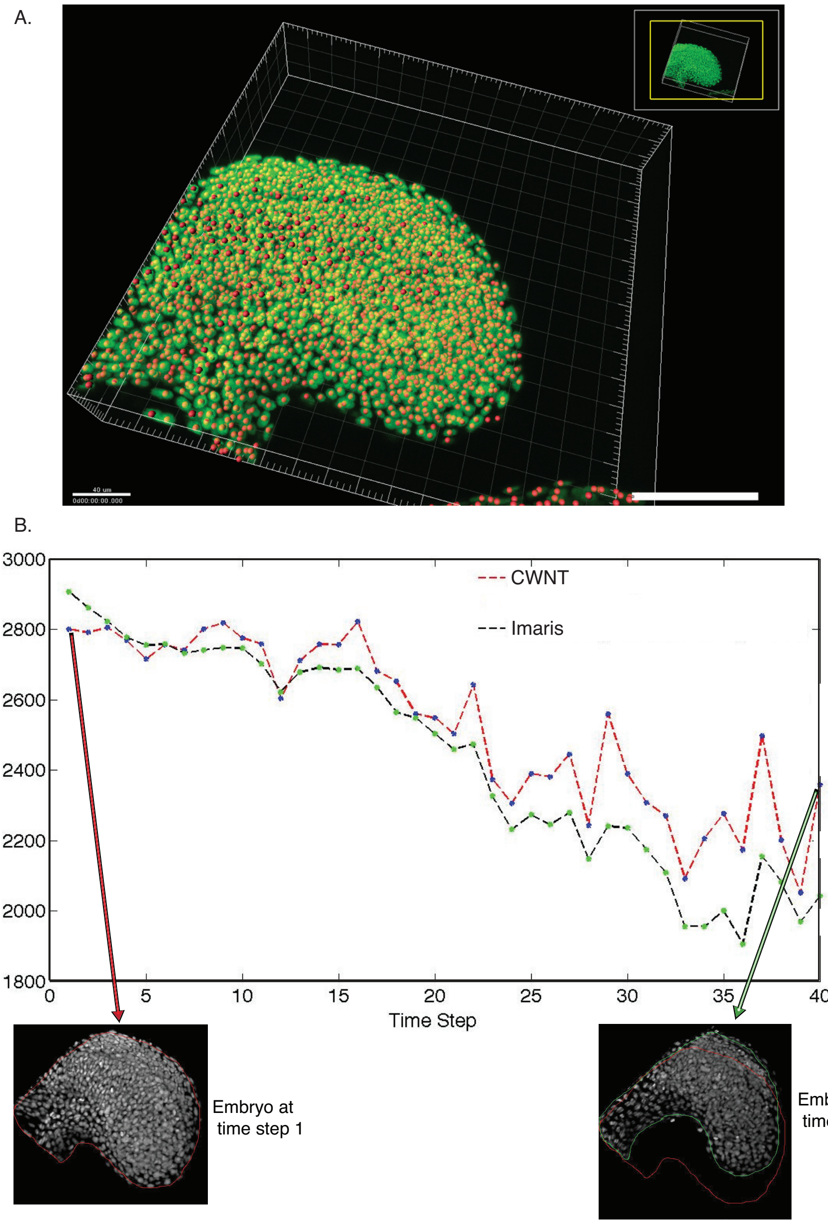


Figure 4.11: A. Snapshot of a single stack of the PSM after nuclei segmentation using a semi-automated method in Imaris. Rendered green nuclei and red centroid spots plotted on top. B. Comparison of total number of nuclei over time with CWNT and Imaris. The number of nuclei decreases over time due to the movement of the embryonic tissue out of the imaging frame owing to development and growth. Embryo outlined (in red) at time frame 1 and in green at time frame 40.

objects. Characteristics of the embryonic image data such as density of nuclei in a volume, signal-to-noise ratio (SNR), background intensity and spatial resolution were utilized to produce synthetic or artificial images as shown in Fig. 4.12A. The advantage of this is that the exact 3D position of every object is known and image qualities such as SNR and density can be continuously varied in 3D. We segmented the objects using the CWNT algorithm and compared the differences between centroid positions after segmentation to the known coordinates as seen in Fig. 4.12B. In order to measure error rates in the algorithm, we defined two quantities, namely sensitivity that represents the detection rate of objects in an image and precision that represents the probability that a detected object is actually a real object, explained in section 6.5.1. The score is given by the product of sensitivity and precision. We examined the efficacy of our algorithm by varying the SNR and the density of objects in the synthetic images.
Fig. 4.12C. shows on the left artificial images for a range of SNR and density values. We segmented the objects for all density values and found that the algorithm is robust for a range of SNR values. Similarly, we tested the algorithm performance by varying density values and found that the performance monotonically decreases with increase in density. Further, we were able to assess that the typical density within the embryonic tissue lies within a range of 0.0022 µm^{-3} to 0.0032 µm^{-3} as highlighted in Fig. 4.12C. right, wherein the algorithm scores about 90%. The plots in Fig. 4.12C. are shown both for a single density and a range of SNR and similarly for a single SNR and a range of densities. The analysis of results of synthetic data was done in collaboration with Dr. Koichiro Uriu.

For a value of SNR=5, which is similar to what we would obtain from our real image datasets, we finally checked the performances of algorithms variants for a range of density values. Fig. 4.12D. shows the dependence of sensitivity and precision on density (section 6.5.1). The performance of the core segmentation algorithm without post-processing (pink open circles in Fig. 4.12D) has high precision whereas the sensitivity drops with increase in density. Combining post-processing methods suggested in this work (k-means and GMM) with the segmentation algorithm improves the performance curves. GMM exhibits superior results than k-means for our artificial data sets. Post-processing with

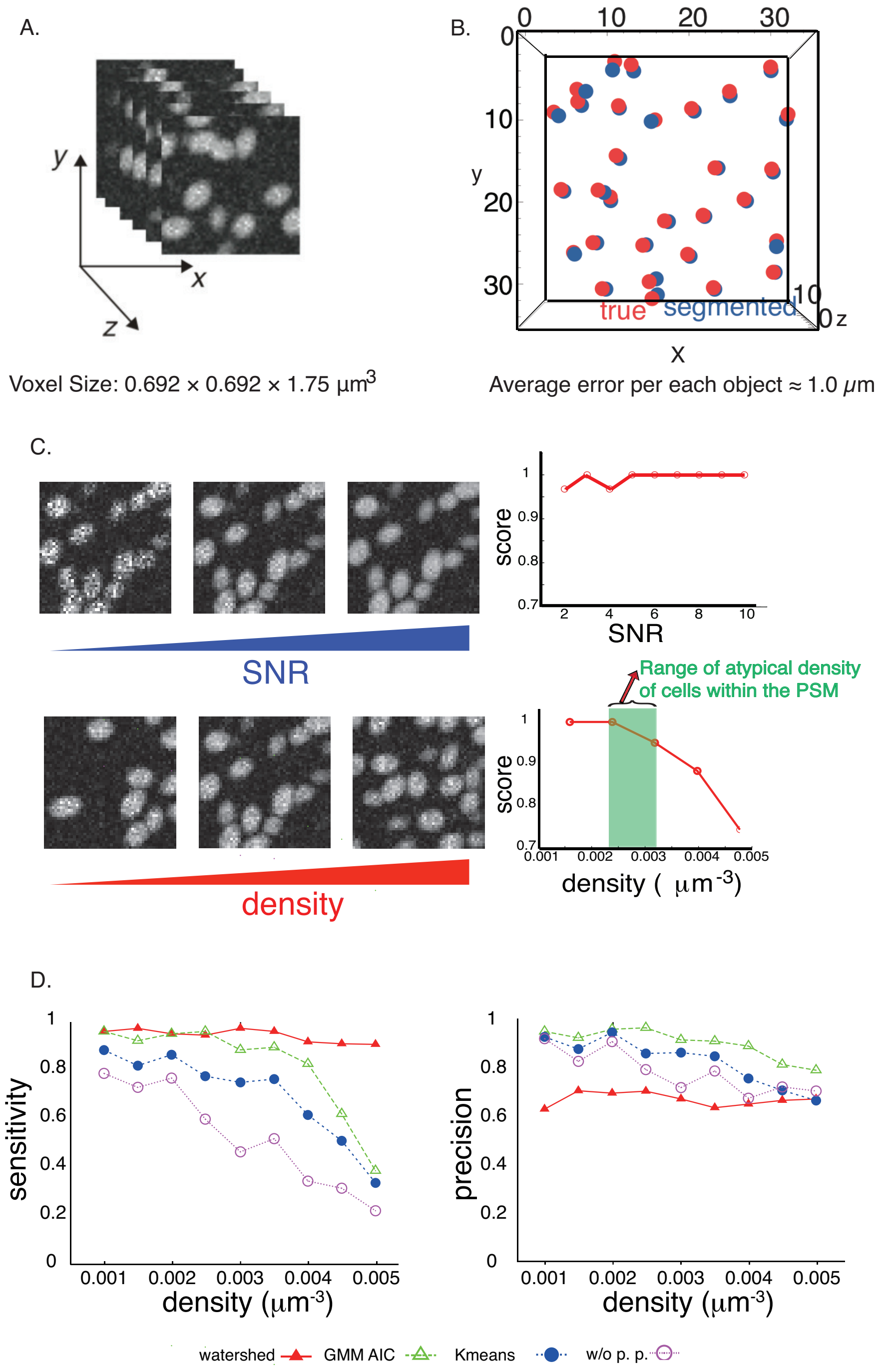


Figure 4.12: A. series of slices from a synthetic image volume. B. Positions of known centroids in synthetic images (blue dots) and centroid position after segmentation (red dots). C. Segmentation algorithm response to SNR and density. Score (given by the product of Sensitivity and Precision) demonstrates robustness of the algorithm with respect to SNR values and a monotonic decrease in performance with increase in density. D. Sensitivity and Precision (measures of false positive and false negative rates) evaluates for different density. We compare the CWNT algorithm without post-processing (pink circles), post-processing with k-means (blue dots) and GMM (green triangle) and finally a commonly used watershed algorithm (filled red triangle).

the standard watershed algorithm shows a bias towards splitting objects (Meyer, 1994) and thereby is able to achieve highest sensitivity (above 90%) regardless of varying density. However, the precision is consistently about 60% even for low densities, indicating that the watershed tends to over segment.
Our analysis with artificial images implies that the efficacy of the algorithm combined with GMM post-processing method is quite robust. The sensitivity is high for density values up to 0.0035 μm^{-3}, which is within the range of densities we have in the embryonic tissues. The precision is of about 90% and is density invariant, indicating that our nuclei segmentation algorithm performs well within the regime of our biological problem of interest. The synthetic data allows a qualitative and quantitative evaluation of our segmentation algorithm and gives us a better understanding of the advantages and shortcomings with respect to testable features of the real world images.

4.6.3 *In vivo* benchmarking of algorithm with embryo chimeras

Synthetic data form an excellent tool to test segmentation algorithms for varying image properties, like SNR and density, however, it is difficult to mimic all the properties of real images artificially. In contrast, in real datasets, owing to the high density, it is hard to benchmark algorithms because the true positions are a priori unknown. Since our motivation to formulate a segmentation algorithm was to segment nuclei in the context of live zebrafish embryo during development, we designed embryo chimeras using transplantation techniques for *in vivo* testing (section 6.6, Fig. 6.2). The experiment supplied natural test datasets within the context of a live embryo to benchmark algorithm efficiency for wide range of nuclear sizes and volumes, image contrast and densities of nuclei within the limits of tissue packing.

Acceptor embryos that had double labeled nuclei within the PSM around the 10 somite stage were screened for live imaging. The embryo was imaged from the 16 somite stage onwards using both channels sequentially on an upright single photon confocal microscope. As a result of the transplantation, the embryos had a fraction of nuclei labeled with a histone fused m-cherry marker, that had corresponding nuclei labeled with histone fused gfp in the other channel as seen

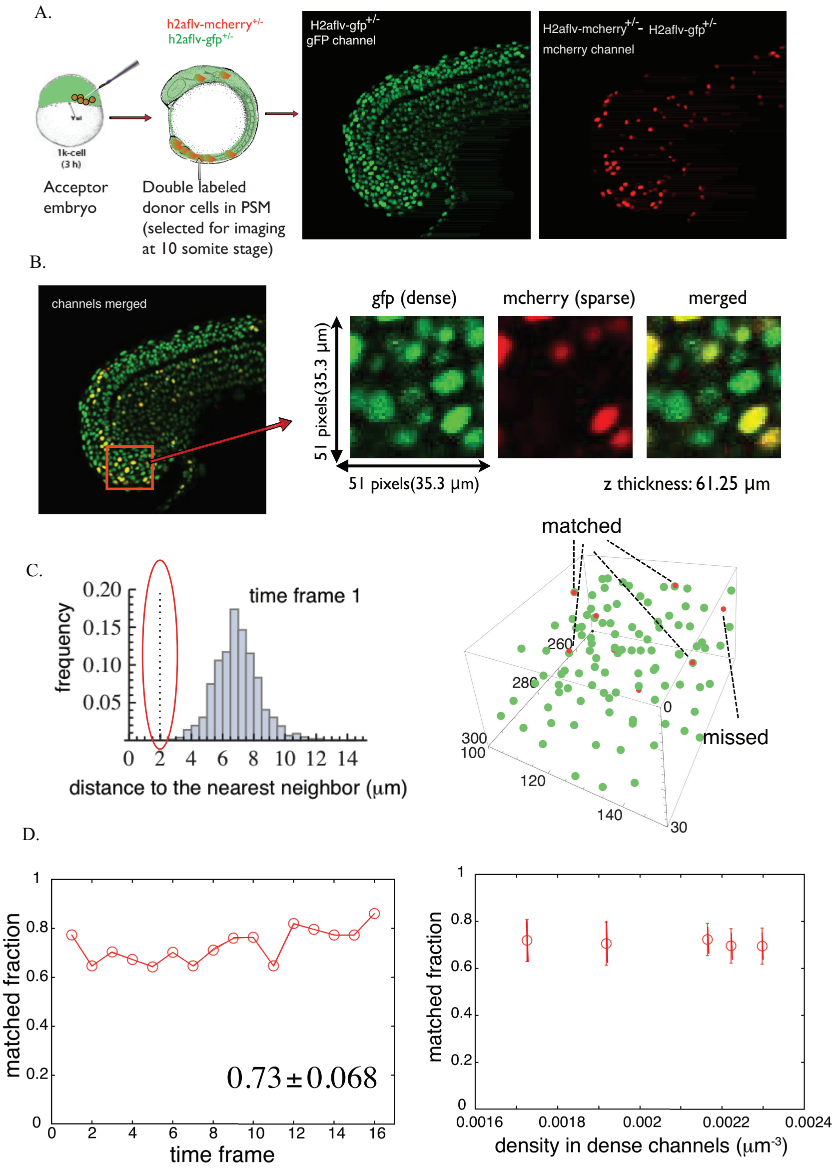


Figure 4.13: A. Doubled labeled cells were transplanted into ubiquitously gfp labeled embryo at 1000-cell stage. Embryos were screened for an appropriate cell density within the PSM in the sparse channel at 10 somite stage. Snapshot of the PSM from live imaging, 16 somite stage onwards for both dense(gfp) and sparse(mcherry) channels, 16-bit grayscale images. B. Images were cubically cropped within the same spatial coordinates in both channels. C. Nearest neighbor distance distribution for all nuclei within a time frame suggest a matching distance cut-off of 2µm. Matches are found if the Euclidean distance between a nucleus of mcherry channel and gfp channel lie within this distance. D. Success rate of the algorithm is measured in terms of matching fraction between nuclei in dense and sparse channels over time. There are, on average 73% matches for a time-lapse movie. Match fraction is found similarly for five different movies with different densities of nuclei in their sparse channel, indicating a consistent match of about $\sim 70\%$.

in Fig. 4.13 A , also see Supplementary movie S5 and the corresponding movie legend in Chapter 7.

Next, small cubes approximately of size 51×51×35 voxels[3], were cropped from different regions of the time-lapse movies as illustrated in Fig. 4.13 B. As a result of cropping, the computational time for nuclei segmentation was fairly short and allowed easy analyses of image properties, such as density and SNR. Nuclei segmentation yielded positions and voxel lists of all nuclei in the dense channel (gfp) and a subset of those in the sparse channel (mcherry). The sparse channel allowed us to achieve reliable segmentation results, since overall density of nuclei compared to the dense channel as well as local density within a volume was quiet low. Due to the lower density, we assured that the positions of nuclei in the sparse channel detected by the algorithm could be considered as "truth". In contrast, we assumed that the segmentation algorithm was more challenging in the high density channel for the same volume. Thereby, using the technique it was possible to evaluate *in vivo* performance of the algorithm. The detected centroid of a correctly segmented nucleus of a transplanted cell in the dense channel image would have a corresponding nuclear position in its close vicinity in the sparse channel (ground truth). Therefore, we counted the number of matches between the two channels and defined a detection rate given by the 'sensitivity', also previously used for the analysis of synthetic images. Centroid matches were found when the position of the nucleus in the dense channel is within 2µm Eucledian distance from a true point in the sparse channel as shown in Fig. 4.13 C.

We defined the match fraction as the ratio of the number of matches between dense and sparse to the total number of nuclei in the sparse channel. Fig. 4.13 D shows the match fraction fairly consistent over time, with an average match fraction of 0.73 ± 0.068. We repeated the analysis in different regions of several embryos with varying densities in both dense and sparse channel and obtained on average match fraction of about 70%, as shown in Fig. 4.13 D (right). Due to large differences in intensity and poor local contrast within real images, we fail to segment nuclei in the dense channel and thereby loose 30% of matches.

The segmentation algorithm combined with benchmarking tools indicate that our algorithm performs well in 3D, and is able to identify objects of different sizes and in different regions of the tissue varying in packing and contrast. In essence, we have converted a 3D image volume from the PSM into quantitative datasets designated with properties such as centroids volume and voxel lists for each segmented object in the image. In conclusion, we have put forth a segmentation algorithm based on image derivatives and have proposed post - processing steps to combine properties of segmented objects to specifically address the problems of under-segmentation, thereby enhance the overall performance. With such encouraging segmentation results, illustrated in Fig. 4.14, we were interested in tracking centroids of segmented nuclei over time.

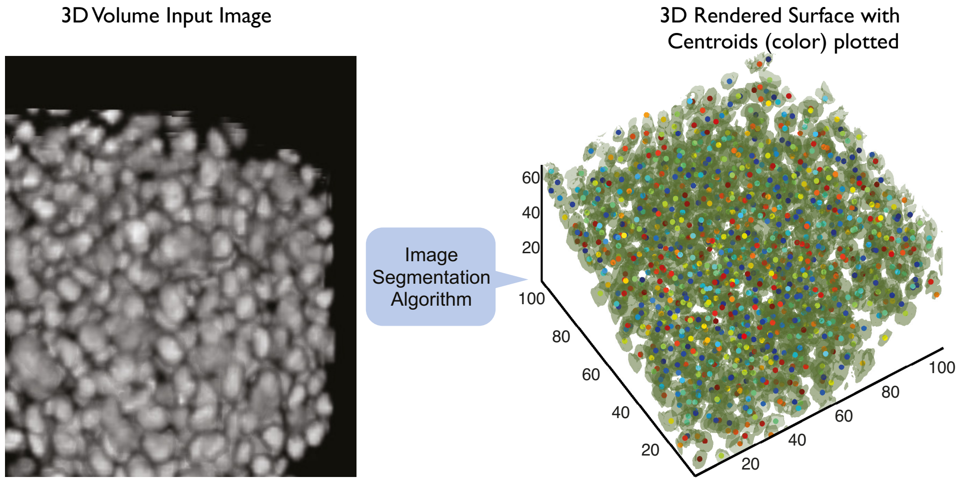


Figure 4.14: Successful 3D segmentation results. 3D input raw image shown on the left and the 3D rendered nuclei (in green), with colored centroids plotted on top after segmentation are shown on the right.

4.7 Nuclei centroid tracking within the PSM

The segmentation algorithm yielded reliable results and it was now possible to track the centroid positions of nuclei over time. In order to precisely track the positions of cells over time, I adapted the algorithm developed by Sbalzarini and co-workers (Sbalzarini and Koumoutsakos, 2005). The Sbalzarini algorithm considers imperfections in nuclei identification and therefore accounts for appearance and disappearance of nuclei in all time frames. It is implemented by formulating an association matrix for all centroids between consecutive time frames. Final nuclei trajectories are then linked by minimizing a cost function that depends on the distances between centroid positions in the subsequent time frames for all time points. We found this to yield good results for our segmented data and serve purpose of our analysis.

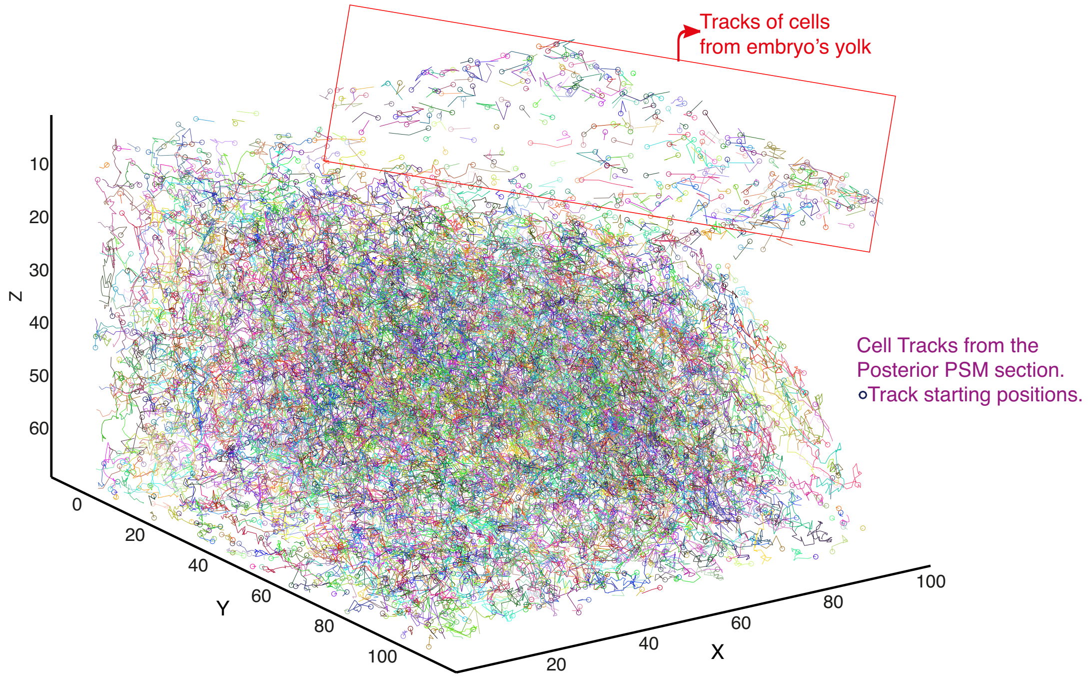


Figure 4.15: Random colored 3D tracks for the posterior section of the PSM of an18 somite stage embryo (70 time points with time interval of 1.4 minutes). Open circles indicate starting positions of tracks.

Fig. 4.15 shows 3D traces of nuclei for a small volume of PSM image. As a result of successful segmentation and implementing the tracking algorithm, it was possible to make reliable trajectory linkages that allows following the 3D centroids of individual nucleus. However, even for this small volume, it is difficult to gauge the tracking performance due to the large number of cells and their tracks. Therefore, we use embryo chimeras for the purpose of inspecting the tracking results.

4.7.1 Trajectory length verification using embryo chimeras

Here, I would like to demonstrate the use of embryo chimeras to better assess the tracking results. The embryos, carriers of gfp and mcherry nuclear markers, formed the ideal test set to compare trajectory lengths in live images of the same embryo, with one of the channels consisting of closely packed nuclei while the other with scattered labeled nuclei, as explained in section 6.6. To demonstrate the usefulness of embryo chimeras to evaluate trajectory lengths, in this section, I would use the embryo shown in Fig. 4.16, which shows a 3D snapshot of a 17 somite stage embryo chimera with gfp (dense) and mcherry (sparse) labeled nuclei (also refer to Supplementary movie S6 and the corresponding movie legend in Chapter 7) .

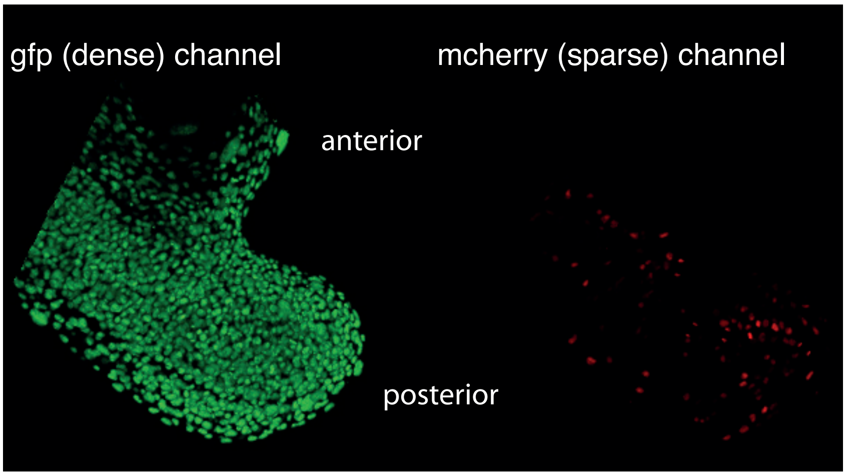


Figure 4.16: Lateral views of the PSM, anterior is left and posterior is right. 3D snapshot of a time step from a 17 somite stage embryo after transplantation experiment. Image on the left shows all nuclei labeled in green, carrier of h2aflv-gfp$^{+/-}$ (dense channel) and on the right, only a fraction of those labeled with m-cherry, carrier of h2aflv-gfp$^{+/-}$ - h2afv-mcherry$^{+/-}$ (sparse channel). Time interval between consecutive time frames is 65.7 seconds with 30 time steps. Voxel size: $0.691 \times 0.691 \times 1.89 \mu m^3$.

The transplant experiments allowed *in vivo* comparison of trajectory lengths in two different scenarios; closely packed nuclei and a subset of those being scattered labeled within the same embryo. The transplanted images were segmented for nuclei in both dense and sparse channel and the tracking algorithm was used to form trajectory links between centroids of nuclei over time. Fig. 4.17 and Fig. 4.18 show the results of inspection of trajectory lengths using embryo chimeras. Please note that the number of trajectories is much larger than the number of segmented nuclei because of large number of broken tracks due to imperfections in segmentation.

The validation using transplant experiments are consistent with the fact that the outcome of segmentation is affected by density (shown earlier in section 4.6.2 and 4.6.3) and therefore also trajectory lengths depend on the density. Fig. 4.17 and Fig. 4.18 illustrate that we can obtain longer tracks for the channel with lower density of nuclei within the PSM when compared to track lengths for the dense channel within the same embryonic tissue over time. In this example, the sparse channel has approximately 6.2% nuclei of the dense channel, as shown in Fig. 4.19A. Fig. 4.19B. further confirms that the fraction of longer tracks are obtained for the sparse channel due to lower density and sparsity of nuclei.

We were yet able to detect the track of the same nucleus coming from the sparse as well as the dense channel. For doing this, a distance cut-off of 2 µm was set to find a match between the first 5 time points of the tracks coming from the dense and the sparse channel. When the distances were within this cut-off, it was possible to trace the tracks of the same nucleus coming from both the channels. Fig. 4.19 C. illustrates plots of 2 nuclei having the same trace, one coming from sparse and the other from the dense channel. Comparison of full length trajectories alone (i.e. trajectory lengths with 30 time points only), gave a 33% match between tracks of dense and sparse channel. Here, for the first time, we provide a glimpse on how embryo chimeras can be useful to compare *in vivo* tracking performances. The next task was to evaluate the dispersion of cells within different regions of the PSM. For this analysis, it became mandatory to have uniform density of cells throughout the PSM to achieve reliable statistics from our data, however it was not necessary to always have longest lengths of trajectories of cell tracks. Therefore, cell track data from dense type of channel would be used for computation of diffusion co-efficient of cells.

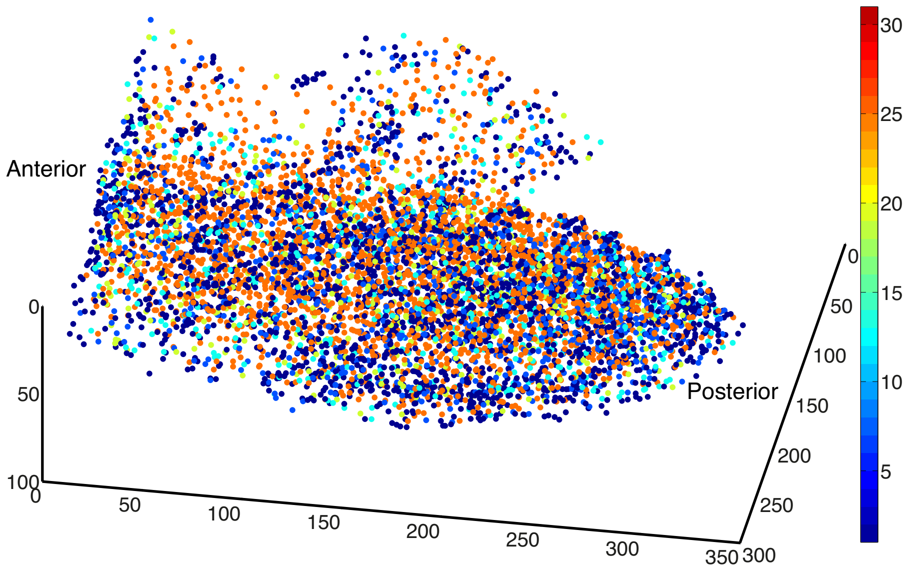


Figure 4.17: **Dense** channel 3D trajectory lengths indicated by the color bar for Fig. 4.16,left. Dark blue color indicates shorter trajectory lengths < 5 and color dots towards red indicates longer trajectory lengths.

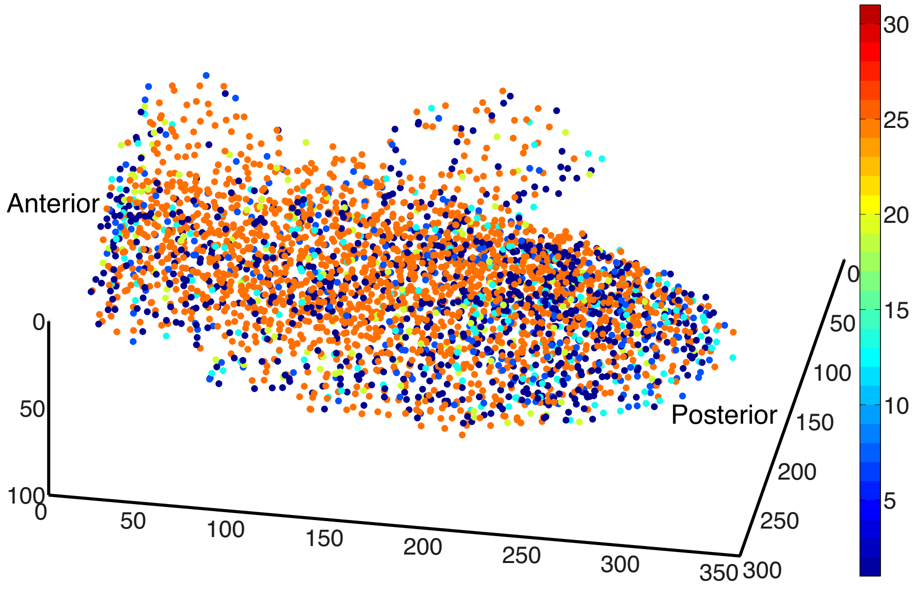


Figure 4.18: **Sparse** channel 3D trajectory lengths indicated by the color bar for Fig. 4.16, right. Dark blue color indicates shorter trajectory lengths < 5 and color dots towards red indicates longer trajectory lengths.

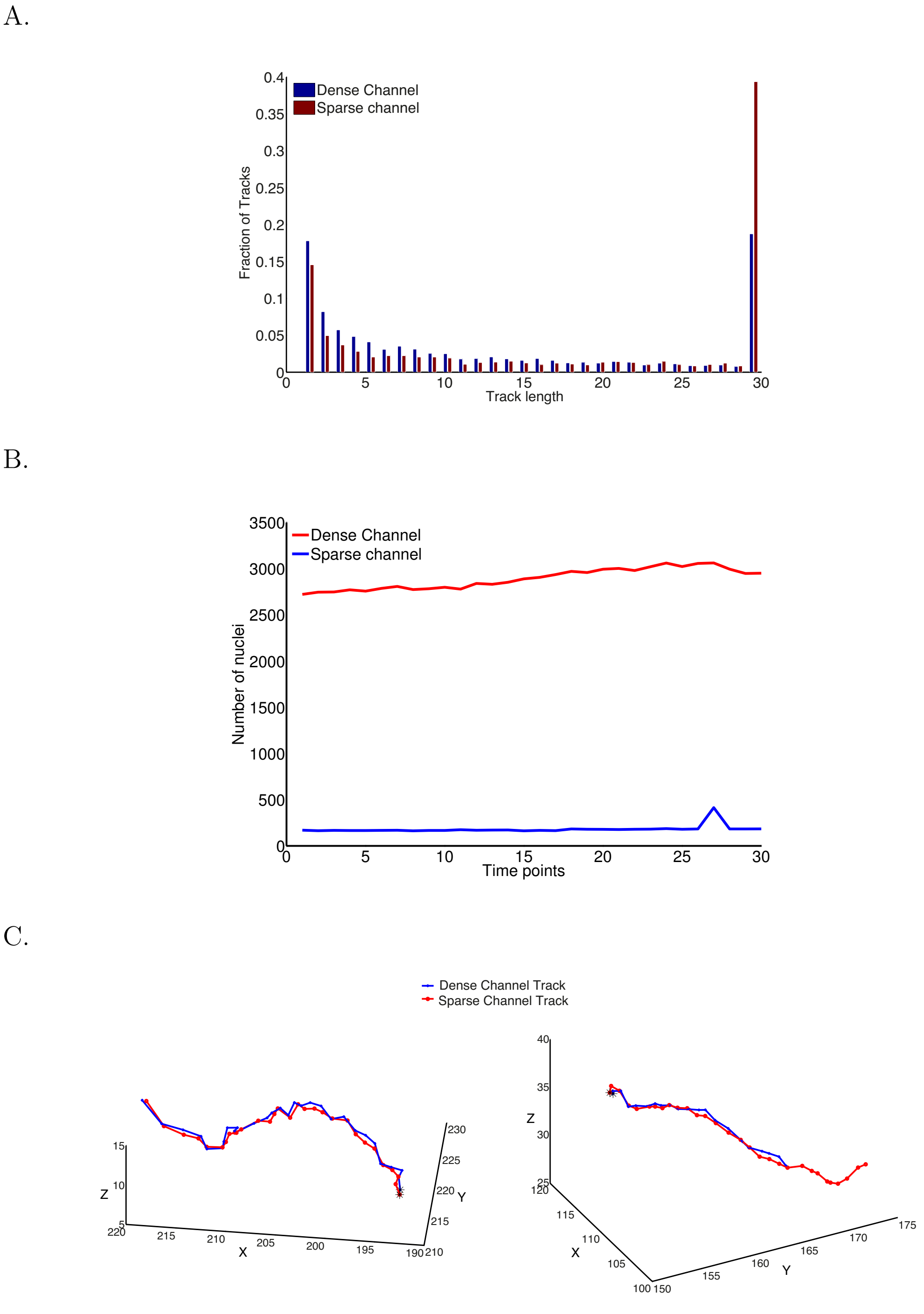


Figure 4.19: A. Trajectory length distribution for both dense and sparse channels. Fraction of tracks with longer trajectories are obtained for the case with lower density of nuclei (mcherry/sparse channel. Due to high density of nuclei, more shorter track lengths are seen for the gfp/dense channel. B. Nuclei number variation for both dense and sparse channel indicated here. On average, the sparse channel has about 6.2% nuclei of the dense channel. C. Examples of centroid track of a single nucleus from both sparse and dense channel. The match between the traces was found within a distance of $2\mu m$. Black asterisk indicates starting positions of the tracks. 33% match of full length track (i.e. upto 30 time points) was found between sparse and dense channel.

4.8 Diffusion co-efficient of PSM cellular oscillators

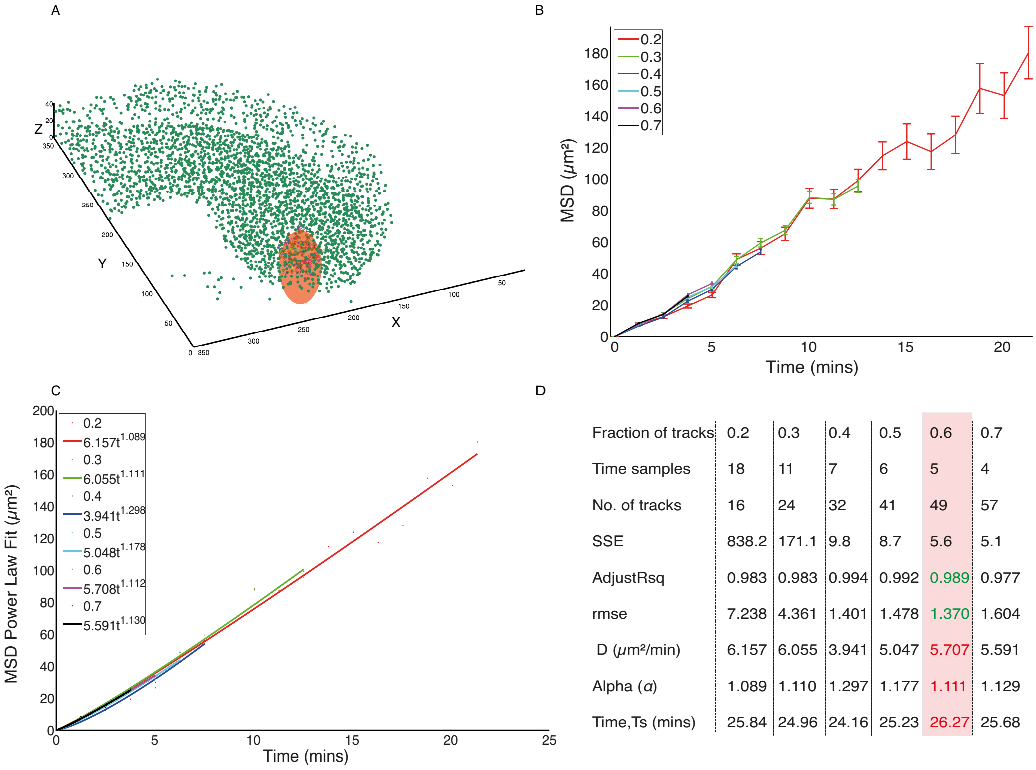
The study of diffusion of cells within the PSM would permit us to evaluate the characteristic time scales of motion of the PSM cellular oscillators. We were interested in inferring the diffusion co-efficient of cells within the posterior PSM where cell mixing is high and is likely random. Further, the diffusion co-efficient can then be re-phrased as a flipping rate of cells for the model (Uriu et al., 2010) discussed in section 6.8. Combining cell tracking data and the model developed by Uriu and co-workers (Uriu et al., 2010), we developed a method to compute the Mean Squared Distance (MSD) for a collection of cells in a fixed spatial region within the PSM and thereby locally infer the diffusion co-efficient of cells (section 6.8). We also evaluated the precision of the estimated diffusion co-efficient of cells by selecting different number of tracks with different trajectory lengths.

The MSD was calculated for a specific region in the PSM tissue using nuclear coordinates of cells within a radius of 18 µm. The method of region selection allowed us to understand time scales of diffusion of cellular oscillators with respect to the PSM reference frame and gain insight on the values along the anterior-posterior axis as well as along the dorso-ventral direction of the embryo. Fig. 4.20 A shows a selected region in the posterior PSM highlighted with an orange ellipsoid for a 19 somite stage developing zebrafish embryo.

In order to obtain a good quantitative description of the motion of oscillators from the data in the posterior PSM, we processed the selected data such that we considered only times up till which a high fraction f_c of the initial tracks exists. The value of f_c controls the selection of tracks (number of tracks and their length) used here to estimate locally the value of MSD and hence we were able to check whether we obtained consistent answers independent of small changes in the selection of tracks. This was done to ensure good statistics for the MSD calculation. Fig. 4.20 B. illustrates the procedures for a selected set of tracks that correspond to values of f_c ranging between 0.2 and 0.7 and shows that all curves are consistent. This plot thus indicates that we would be able to get a reliable estimate of diffusion coefficient of cells (D) and that artifacts arising from imperfections in segmentation and tracking would not significantly affect our estimated value from the data. Next, for each MSD curve, we used a power law fit, $MSD = Dt^{\alpha}$ (see section 6.8) to estimate the values of D, α and their corresponding time scales T_s, as illustrated in Fig. 4.20 C. The plot shows for each f_c value that ranges between 0.2 and 0.7 that we can obtain D and α, given by equation (6.19) in section 6.8. Therefore, to estimate the best fit, we used the goodness-of-fit statistics and evaluated sum of square of error (sse), adjusted-R-square and root mean squared error (rmse) and chose that value of D, α and T_s, which was corresponding to least error, highlighted column in Fig. 4.20 D. The number of tracks and time samples vary depending on f_c, as does the quality of fit and the final values of D, α and T_s.

For the 19 somite stage wild type embryo, imaged at 28°C, shown in Fig. 4.20, the estimated values were, $D = 5.7$, with $\alpha = 1.1$ and therefore, the time scales of motion of oscillators was found to be $T_s = 26.2$ minutes. The value of α indicates that there exists a super-diffusive process (since $\alpha > 1$) in the posterior PSM. Similar analysis for a *Trilobite* mutant (convergence and extension phenotype) at the 19 somite stage at 28 °C revealed $T_s = 31.3$ minutes, $D = 4.5$ and $\alpha = 1.1$ (analysis details are not shown here, see Supplementary movies: S7-A and S7-B with movie legend in Chapter 7). The results indicate that the diffusion of cells in the *Trilobite* mutant is slower in the posterior PSM than their WT counterparts. In order to have a broader overview of diffusion of cells in the PSM, we would be able to estimate T_s in WT and mutant embryos for different developmental stages and temperatures in the near future.

Next, we wanted to estimate D and T_s throughout the entire PSM (for the WT embryo shown in Fig. 4.20) by considering smaller sub-regions along the anterior-posterior (AP) axis as well as dorso-ventral axis. Therefore, the present method was modified to compute D, α and T_s for sliced sections along the z direction at intervals of about 10 µm as shown in Fig. 4.21 B. This enabled us to have a pictorial view of how the time scales change along the AP axis, and as well as along the lateral (or ventral to dorsal) direction of the tissue, as we move from the PSM cells close to the epithelial until the cells of the PSM spatially closer to the notochord are reached. Fig. 4.21 A shows a 3D view (obtained with Fiji 3D viewer) of the 19 somite staged embryo imaged at time interval of 75.3 seconds used here for the purpose of demonstrating the computation of diffusion coefficient of cells for the entire PSM tissue. Fig. 4.21 B illustrates 4 sections along



Fraction of tracks	0.2	0.3	0.4	0.5	0.6	0.7
Time samples	18	11	7	6	5	4
No. of tracks	16	24	32	41	49	57
SSE	838.2	171.1	9.8	8.7	5.6	5.1
AdjustRsq	0.983	0.983	0.994	0.992	0.989	0.977
rmse	7.238	4.361	1.401	1.478	1.370	1.604
D (μm²/min)	6.157	6.055	3.941	5.047	5.707	5.591
Alpha (α)	1.089	1.110	1.297	1.177	1.111	1.129
Time,Ts (mins)	25.84	24.96	24.16	25.23	26.27	25.68

Figure 4.20: A. Centroid plot after segmentation (green dots) for first time point . Tracks from the posterior PSM of the embryo are selected in a 19 somite stage embryo imaged at 28°C, at a time interval of 75.3 seconds, highlighted by the orange ellipsoid. B. Dependence of MSD results on f_c. Linearity of the MSD curves for different sets of selected tracks within the same posterior PSM region consistently indicates that the movement is random. The selection of tracks with different lengths does not affect the estimate of MSD. C. Power law fit (given by $MSD = Dt^{\alpha}$) used to determine diffusion co-efficient of cells (D), α and T_s within the selected region for different set of selected tracks based on f_c in the posterior PSM. D. Tabulated overview of the effect of varying the fractional cut-off (f_c) resulting in variation of the number of time samples and total time points. Power law fit is evaluated for each MSD curve and the values of D, α and T_s are chosen for that set of tracks with least error (using goodness-of-fit statistics), which is the highlighted column.

the z-direction, 10 µm apart within the PSM in the dorso-ventral direction. The diffusion values were estimated around each of the sections in the PSM within a 3D radius of 18 µm for all cells in order to estimate the values of time scales (T_s) of diffusion along the AP axis and in the lateral direction over the entire PSM space.

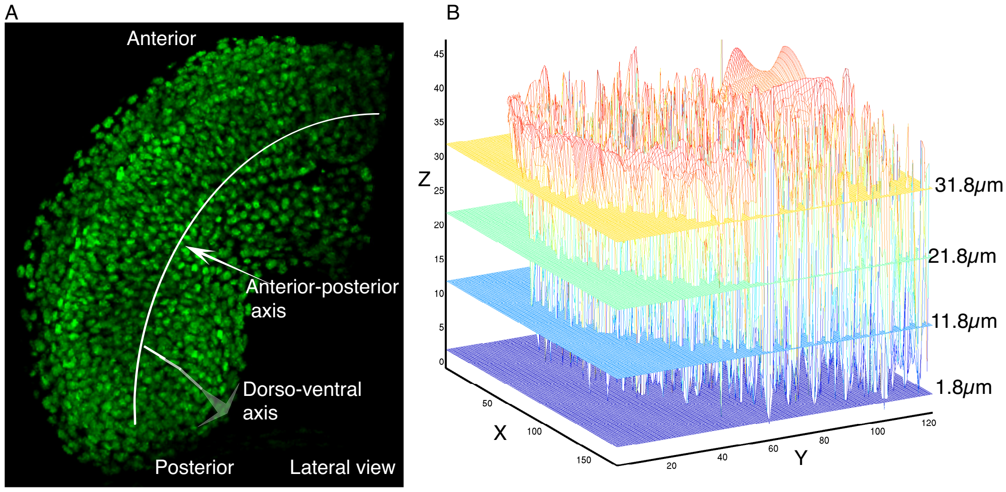


Figure 4.21: A. 3D view of a 19 somite stage WT embryo tagged with histone-gfp, imaged at 28°C at time interval of 75.3 seconds. White curved line indicates anterior-posterior axis and arrowed line shows the dorso-ventral direction of the PSM. B. Mesh plot of 3D centroids a single time point with 4 z-slices that section the PSM along the z-direction into 2d-planes (surface height of Z is indicated by the color). The diffusion coefficient of cells is estimated in the z-sections spaced by 10 µm.

As described, we obtained T_s values for cells around each of the 4 sliced sections within a radius of 18 µm for the entire image, as seen in plots Fig. 4.22 A-D. The plots encapsulate T_s values for the entire 3D image in 4 different lateral sections of the embryo, using a linear colormap index, indicated by the color bar. Similar plots were also made for D and α (data not shown here). From the plots in Fig. 4.22, it can be inferred that the oscillators diffuse faster in the posterior PSM and relatively slower as we move towards the anterior PSM. Colorbar on the plots determine that the time scales of diffusion of cells in the posterior PSM is about 25-30 mins and as we move towards the anterior, the diffusion time scales ranges between 55-60 mins. As the 4 plots show similar distribution of T_s for the entire PSM tissue, independent of which sliced section of the PSM they are centered around, we may conclude that the diffusion does not vary strongly along the lateral direction. In other words, the time scales of diffusion are fairly similar along the lateral direction as seen from the plots in Fig. 4.22 A-D.

In summary, the analysis suggests that the diffusion is fairly homogeneous in the dorso-ventral direction of the PSM, whereas the time scales of diffusion increases as we move from posterior to anterior as observed in all the plots. The methods developed here may be further used to systematically analyze the diffusion of cellular oscillators at different stages of embryo and at different temperatures in more detail in the near future. We would also be able to use drugs like blebestatin in WT and mutant embryos and infer variations in diffusion co-efficient values.

To further our understanding of the relationship between movement and synchrony, in the next section, I will briefly outline the method we have established through this work to extract trajectories of *in vivo* single oscillators at cellular resolution.

4.9 *her1* cellular oscillations within the PSM

Previous studies have established that the Delta-Notch coupling is mandatory for keeping the autonomous cellular oscillators, *her1* and *her7*, in a synchronized state and that perturbation to the pathway leads to erroneous boundary segments, also discussed in chapter 2, section 2.1. These studies used in situ hybridization techniques on fixed embryos to analyze segmented boundary formation. They infer coupling strength, defect rates and synchronization response times to perturbations due to addition of drugs, such as DAPT, and recovery rates after effacing drug effects from the embryo (Riedel-Kruse et al., 2007; Jiang et al., 2000). The tools developed through this work would make it possible to read *in vivo* oscillations within the PSM at cellular resolution and therefore, we would be able measure the differential cellular patterns during perturbations as well as recovery from perturbations for the cellular oscillators in space over time. Previously in section 2.1, we illustrated by Fig. 2.5, that the tissue level protein oscillations of *her1*-yfp that occur autonomously in PSM cells are synchronized with adjacent cells, therefore a tissue-level wave-like gene expression pattern is seen sweeping the PSM periodically concomitant with somite formation. Fig. 4.23

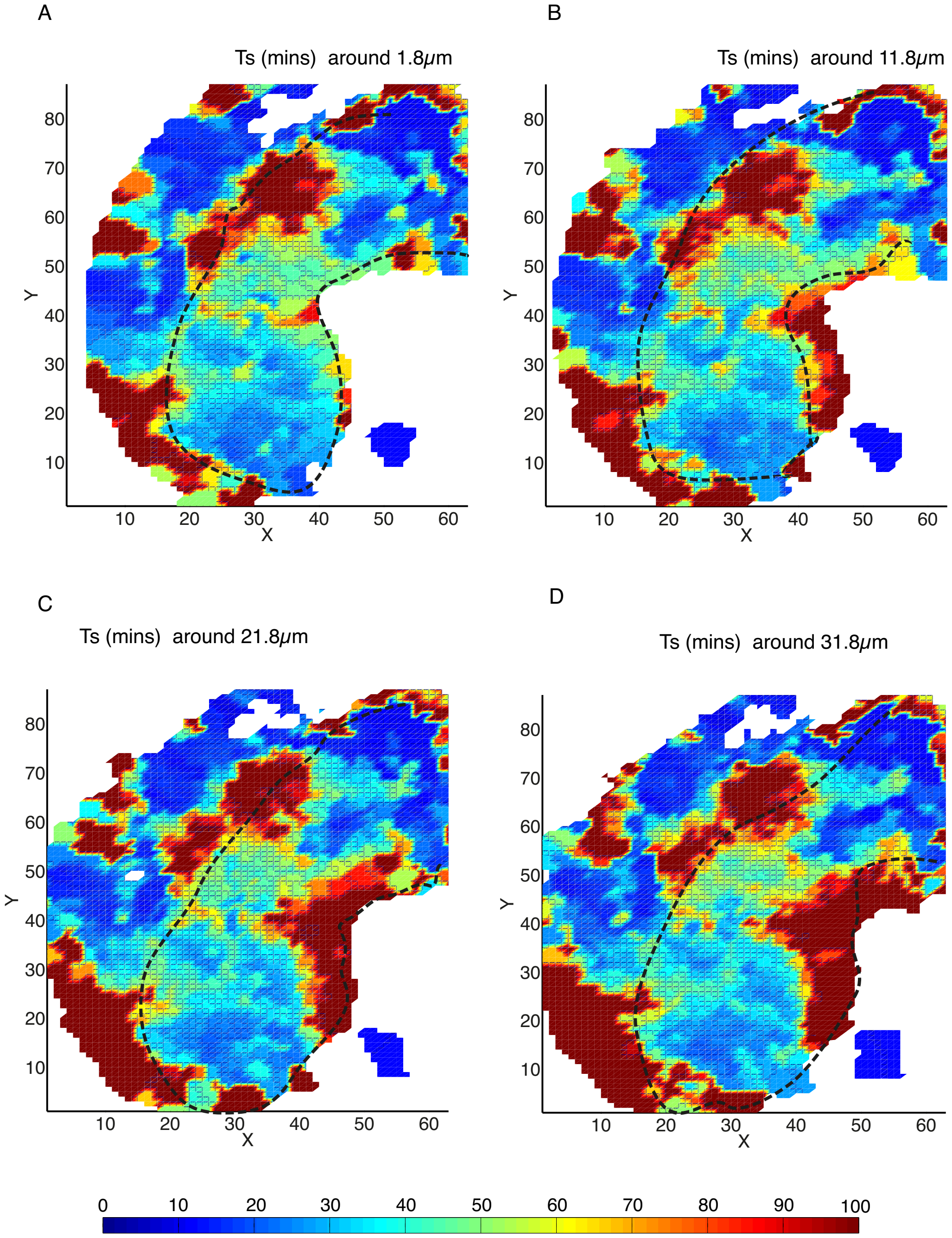


Figure 4.22: T_s plotted for the 19 somite stage embryo shown in Fig. 4.21A. Regular grid of 3D points on the 4 sliced planes was generated as shown in Fig. 4.21B. For each z-section, T_s was calculated for the tracks starting within a 3D radius of 18µm around each cell closest to 3D grid location. Black dotted lines approximately border the PSM tissue in the plots (sketched by hand on the plots). T_s for posterior PSM ranges between 25-30 mins and as we move towards anterior the time scales change to about 55-60 mins, as seen using the color bar.

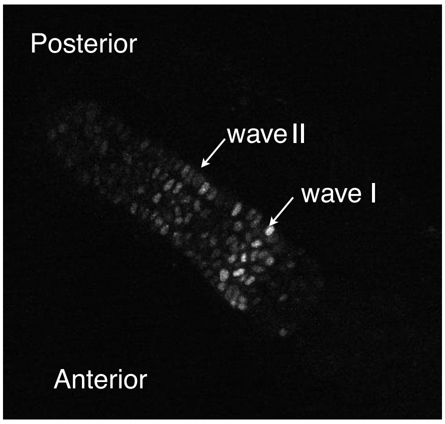


Figure 4.23: Phases of cellular oscillators are in synchrony as they traverse the tissue in the form of a stripe from posterior to anterior. Here, two *her1* stripes are seen in their on-state phase of oscillations. *her1* oscillates autonomously in single cells, however because of delta-notch coupling, the neighboring cell phases are synchronized and their *her1* protein expression traverses the cells from posterior to anterior in the form of a wave until the last formed somite.

shows a snapshot from live imaging of the same *her1*-yfp transgenic line at cellular resolution, wherein phases of cells are seen synchronized as they traverse the tissue from posterior to anterior of the PSM. The figure highlights two waves of expression where neighboring populations of cells are in their "on-phase" of oscillation, also seen a movie on a 2D plane of the *her1*-yfp in Supplementary movie: S8-A with movie legend in Chapter 7.

However, in order to track down single cell oscillators of the PSM, we needed a ubiquitous marker as a reference that would allow us to follow nuclei over time and be able to read their *her1* oscillations over time. A similar protocol has been established previously to read dynamics of an oscillator reporter gene (Swinburne et al., 2008). Here, we used transgenic lines that were carriers of h2aflv-mcherry that would be used to track all the nuclei and read the phases of *her1*-yfp PSM cellular oscillators as seen in Fig. 4.24 (Supplementary movie: S8-B with movie legend in Chapter 7). From these movies and the algorithms discussed in this work, it was possible to segment and track all the nuclei of the mcherry channel. It was also possible to record the voxel lists of segmented nuclei over time which enabled us to read the voxels of the *her1* channel of the same segmented nuclei. Fig. 4.25 shows the average nuclear intensity of the same cell in both the mcherry

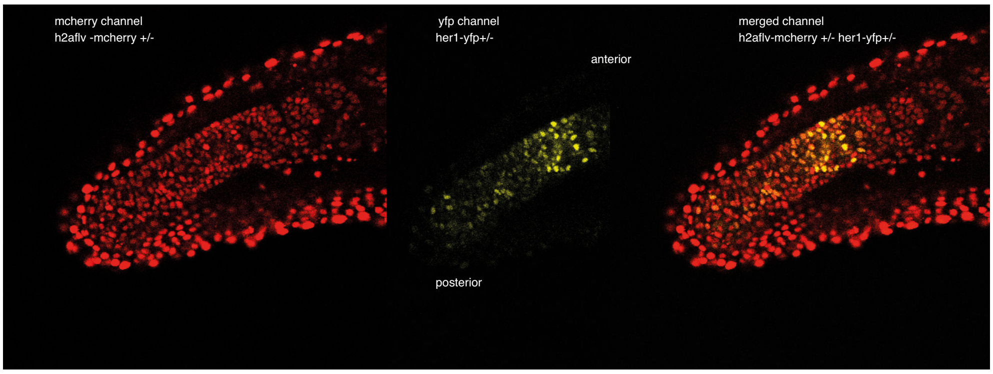


Figure 4.24: 25 somite stage embryo labeled with h2aflv-mcherry and *her1*-yfp imaged at 28°C on a Zeiss upright single-photon microscope. Image taken sequentially at time interval of 89 seconds. m-cherry nuclear marker enables segmentation and tracking of all nuclei and as a result nuclei centroids and their corresponding voxel list are obtained from this experiment.

channel and the yfp channel. Using this method, it is therefore possible to capture *in vivo her1* oscillations of single cells over time. The cherry signal is stable in all the 4 plots indicating no cross-talk between channels, while *her1* signal intensity varies depending on the phase of the cellular oscillator. However, using the present confocal technique with limited imaging hours, it was not possible to obtain very long time lines of oscillations. In order to precisely read phases of oscillators for at least more than 5-6 cycles, we hope to improve on imaging techniques that would enable us to record images for several hours without issues of photo-toxicity in future.

Based on these initial results, it will be possible in the future to deduce synchronization index parameter for a population of cellular oscillators in the PSM reference space and its evolution in time. We can also derive several other parameters like the correlation index, or order parameter that would cast light on synchrony dynamics of cellular oscillators. Such results are not included in this thesis, however, the present framework provides a foundation for measuring them. This will allow us to better understand the interplay and dynamics of synchrony of oscillators, their coupling strength, intrinsic noise within the oscillators and how these quantities are affected by perturbations.

Using the tools described here, we would be able to analyze the synchrony dynamics of oscillators in the *Trilobite* mutant background embryos, where we

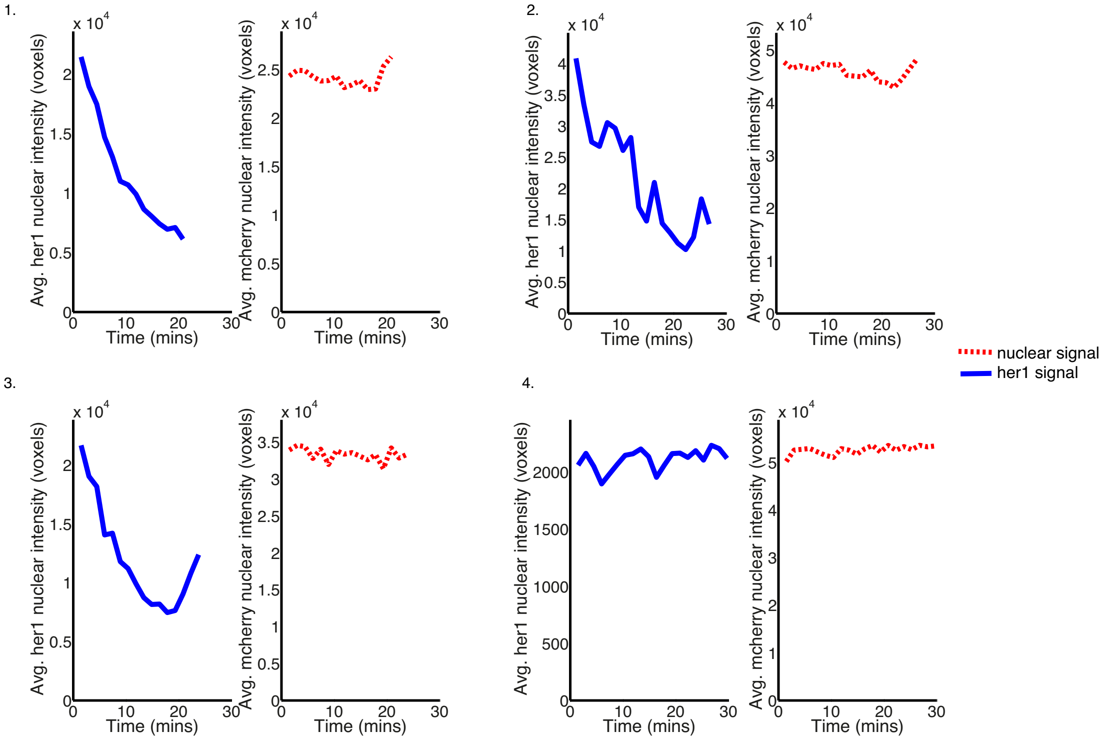


Figure 4.25: Left panels are posterior and right panels are anterior. Average intensity of nuclear voxels as a function of time for 4 different cells with *her1-yfp* on the left (blue) and mcherry on the right (red). Stable nuclear intensities are seen for all the 4 nuclei for the m-cherry channel. The *her1* signal intensity varies depending on the cell's phase of oscillation. Plots 1-3 show oscillatory behavior of the *her1* signal whereas nearly constant nuclear intensity in plot 4 on the bottom right, indicates that the cell does not oscillate.

have already established that the diffusion of cells in the posterior PSM is slower compared to their WT counterparts. Fig. 4.26 shows two waves of *her1* oscillations on a Trilobite mutant background with shorter inter-stripe distance between the waves (Supplementary movie: S9-A and S9-B with movie legend in Chapter 7.

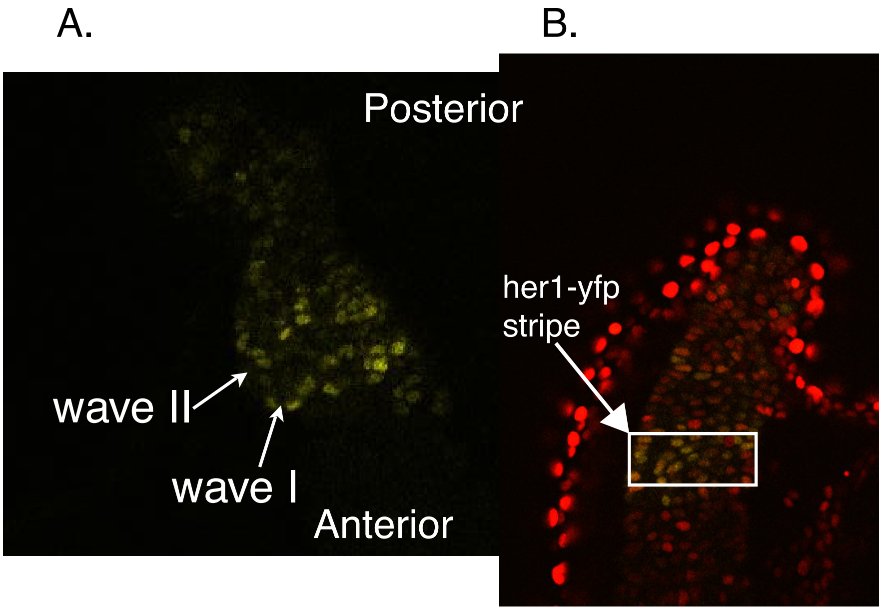


Figure 4.26: A. *her1* oscillations seen on a *Trilobite* mutant background, 17 somite stage embryo imaged at 28°C. Size of *her1* stripe and the inter-stripe distance is smaller compared to WT concomitant with somite formation in *Trilobite* mutants. B. Snapshot from a image file showing *her1*-yfp oscillations and h2aflv-mcherry labeling all nuclei in the *Trilobite* mutant embryo imaged at 28°C.

Chapter 5

Discussion and Outlook

In this thesis, I have developed a protocol that achieves high quality images of the developing zebrafish PSM tissue for quantitative analysis using image processing tools in order to specifically address questions regarding movement of cells and their relation to synchronization properties in the context of the segmentation clock. We used fluorescently labeled nuclei as readouts for cell positions and obtained good images qualified to determine precise positions of nuclei over time. To obtain positions of nuclei, a segmentation pipeline was developed in this work that was carefully constructed to segment 3D objects of about $\sim$2000 closely packed nuclei within an embryonic tissue (PSM) undergoing morphological changes during development. The intention was to provide an algorithm that is well balanced in terms of accuracy, performance and speed. Accuracy of segmentation performance is important to obtain long traces of nuclei that are later used to compute diffusion co-efficient of cellular oscillators and extract their phases over time during development. Speed of algorithm became vital since we wanted to analyze embryos (WT and mutants) at different developmental stages and at different temperatures.

The image processing tools developed through this work have the ability to handle multiple touching nuclei, moving at different velocity, that may divide, enter or leave the observation volume in the growing pre-somitic mesoderm. We provided quantitative validation using artificial data (provided by Dr. Koichiro Uriu, post-doc from the group) to ascertain the efficacy of segmentation algorithm with respect to testable features of images such as SNR and object density and further quantify detection success to demonstrate the merits of the algorithm. For the first time, we substantiated the consistency and reliability of the algorithm's performance within the PSM reference frame under different nuclear density conditions via an insightful tool used in developmental biology, i.e. analysis of chimeric embryos. Cells from a transgenic embryo with double-labeled ubiquitous nuclear marker (h2aflv-mcherry and h2aflv-gfp) were transplanted into another transgenic embryo labeled with only h2aflv-gfp marker (transplantations by Dr. Andrew Oates, supervisor and Dr. Guillaume Valentin, post-doc from the group). This resulted in embryos that had a small fraction of nuclei labeled with both fluorophores, and all nuclei labeled with another fluorophore. Thus the experiment gave a natural environment, which could be usefully tested to assess the segmentation algorithm performance in the context of the nuclear density within the embryo. We were hence able to compare the algorithm performance to identify nuclei in sparsely scattered channel and the same ones in the densely packed clusters within the same tissue. Hence, we demonstrated the usefulness of our image segmentation tool with the help of artificial datasets and *in vivo* benchmarking techniques that aid towards confident quantitative analysis.

With good success with image segmentation, we modified and used the Sbalzarini tracking algorithm (Sbalzarini and Koumoutsakos, 2005) that considers segmentation imperfections by accounting for appearing and disappearing nuclei in all time frames. The algorithm makes trajectory links by minimizing a cost function that is formulated using the Euclidean distance between centroid positions for all time frames. The implementation of cost function is an improvement on the basic nearest neighbor distance, since it considers trajectory links by optimizing the distance between consecutive time points over all time frames. In order to ascertain the degree of tracking performance, we again used embryo chimeras. With the segmentation algorithm, we already established that it was easier to segment the sparse channel successfully and thereby it was also easier to obtain longer fraction of tracks for the same. However, we were still able to find a number tracks of the same nucleus from both dense and the sparse channel that were in good agreement.

From here on, the tools enabled us to quantify a plethora of properties of the system at cellular resolution to further our understanding about the dynamics of cell movements during segmentation in the PSM, and how this affects the synchronization of the oscillating cells.

We started by experimental quantification of movements and relate the data to moving rate parameter as described in Uriu et al. (2010). This work was done in close collaboration with Dr. Uriu who developed a theoretical model for cell movement and their mobility properties with respect to synchronization of oscillations with the PSM of the zebrafish embryo. We developed a method to measure the diffusion coefficient of cells within the PSM tissue of the embryo along the anterior-posterior axis in WT and mutants. Our analysis suggests that the cells diffuse faster in the posterior PSM and slower in the anterior PSM. Previously, similar observations have been made within the chick embryo PSM, although the quantitative analysis was implemented differently (Bénazéraf et al., 2010). We further also deduced from our analysis that the diffusion along the dorso-ventral axis does not vary strongly within the PSM. The method developed in this work would allow us to gain more insight on this parameter of the PSM and thereby gives future scope to analyze WT embryos at different stages as well as mutants in order to better understand the diffusion of cells within the PSM.

The tracking data also enables to capture phases of *in vivo* oscillators that would provide huge scope to verify time series correlations of oscillators in different region of the tissue over time. This would allow us to visualize and quantify the cellular level details for better understanding of the clock dynamics. It would also allow us to compute change in phases of cellular oscillators in WT and mutant embryos, responses to perturbations to drugs such as DAPT. Additionally, we would be able to use this information to extract the synchronization index (Uriu et al., 2010), synchrony parameters such as, coupling strength and compute quantities such as spatio-temporal correlation functions from the real data.

Although, at this point it would be difficult to give a simple straight answer as to *how cell movements affect cell synchronization properties*, but we have gained considerable insight through this work on time scales of motion of cellular oscillators and developed the tools that would enable us to tweak several properties of the system to further our understanding of the behavior of cellular oscillator and how their assembly may generate a precise molecular segmentation clock.

It is suggested that diffusion of oscillators is an inherent feature of the PSM and may be linked to the age and the dimensions (owed to continuous tissue level convergence and extension movements) of the PSM. Further, the diffusion slows down as we move towards the anterior PSM. We know that synchronous oscillations emerge at the posterior PSM of the tissue where cell mixing is relatively high, and the oscillations slow down as we move towards the anterior as well, and eventually are arrested in the form of somites. On the other hand, experimental and theoretical studies on the clock so far have emphasized on the role of coupling of cellular oscillators as an essential machinery for proper functioning of the segmentation clock. Several studies have simulated inter-cellular coupling of phase oscillators Cinquin (2007); Herrgen et al. (2010); Morelli et al. (2009); Uriu et al. (2012) and have considered the coupling strength to be fairly constant throughout the PSM. However, I would put forward a hypothesis where we could consider a space-varying coupling strength along the posterior-anterior axis that changes relative to diffusion of oscillators and therefore may affect their synchrony dynamics properties. It is further speculated that the ratio of time scales of diffusion of oscillators and the coupling strength (a measure of local synchrony among a population of oscillators) within a spatial region of the PSM maintains an inter-relation such that a tissue level output is achieved whose products are waves of oscillating gene expression and dynamic re-arrangement of cell that ultimately form somitic furrows.

Advances in microscopy techniques and image processing tools provide great opportunity to quantitatively evaluate and characterize cellular and sub-cellular level dynamics of processes such as collective motion, gene expression patterns within the developing embryo. Thus it is increasingly becoming vital to provide computational tools that have the ability to convert high resolution image data into reliable quantitative measures in field of developmental biology. They may further bridge experimental observations and theoretical models for better understanding of how certain processes are regulated and coordinated to form an embryo. Here, we have optimized live imaging and image processing tools to suit reliable quantification of the PSM of the zebrafish embryo. This further helps us to address specific questions pertaining to oscillators motion and their dynamics at cellular resolution and allows interplay of mathematical models to interpret experimental data.

Through imaging techniques and tools, it would be possible to provide a better quantitative description of such parameters of the clock and propose mathematical models that would include experimental observations. For instance, quantitative measures of the diffusion of oscillators may be used as an input to existing theoretical model (Uriu et al., 2010), and thereby provide the model with better predictive power for segmentation clock. Further considerations to other structures like filopodia, actin and myosin networks can be examined to understand the their contribution to the dynamic re-arrangement of cells and their subsequent consequences to the synchronization properties of the clock.

In future, it should also be possible to use new microscopes such as the SPIM (Supplementary movie: S10 with movie legend in Chapter 7), that has the ability to image at a higher frame rate with better spatial and temporal resolution without issues of photo-toxicity and bleaching. This would result in better image segmentation results and enable us to have longer tracks of cells over time and even over several somites. This may enable us to manipulate molecular level features of the clock, and help us investigate cellular level properties such as as motion and synchrony, and how these affect tissue level properties of the segmentation clock. In future, all this together may throw some light on the design principles of the segmentation clock and how different aspects of the clock are linked with each other.

Chapter 6

Materials and Methods

6.1 Materials

- E3 Composition:
    - 5 mM NaCl
    - 0.17 mM KCl
    - 0.33 mM CaCl2
    - 0.33 mM MgSO4
    - 10-5% methylene blue
- Molds for time-lapse imaging (Herrgen et al., 2009)
- ethyl-*m*-aminobenzoate methanesulphonate (Tricane)
- Danieau's
    - 0.4 mM MgSO4
    - 0.6 mM CaCl2
    - 0.7 mM KCl
    - 58 mM NaCl
    - 5 mM Hepes pH 7.6

6.2 Fish care and mutant stocks

Zebrafish Danio rerio were raised and kept under standard laboratory conditions (Westerfield, 2000). Embryos obtained from natural spawnings were staged as described in Kimmel et al., 1995. All wt fish used in this study were of the AB and TL strain. Mutant alleles used here were trilobite (Jessen and Solnica-Krezel, 2004; Jessen et al., 2002), knypeg (Topczewski et al., 2001) and after eight (van Eeden et al., 1996; Jiang et al., 1996).

6.3 Sample Preparation and 4D Image Acquisition

We carried out confocal live imaging of the zebrafish embryos, carriers of h2aflv-gfp or h2aflv-mcherry or textither1 oscillating protein expression in different experiments including wt and mutant backgrounds. From about 16 somite stage onwards, the dechorinated embryos were positioned laterally in grooves that fit embryo head head, made of low melting point-agarose (1.5% LMP agarose with E3 (without methylene blue) and 0.02% ethyl-m-aminobenzoate methanesulphonate (Tricane),) in a petri-dish (Herrgen et al., 2009; Schröter et al., 2008). The dish was filled with E3 (embryo medium) and Tricane, that prevents embryos from twitching during imaging. An upright Zeiss confocal microscope with 40x/1.0 NA water objective and 488nm/ 561nm/ 514nm laser light excitation was used (depending on the flurophore), enabled us to obtain several good quality 2D images (8-bit or 16-bit grayscale) of the nuclei parallel to the xy focal plane within the PSM tissue. Several z-slices (with overlap) of the PSM tissue were made and time series of the images with best possible quality were recorded for 2 to 2.5 hours as shown in Fig. 6.1. We typically imaged 512 × 512 in x-y and 20 - 30 slices in z respectively, with voxel size of $0.691 \times 0.691 \times 1.75\mu m^3$. Imaging experiments were carried out between 23°C and 28°C. For developing segmentation algorithm of nuclei, no temperature controls were done, and hence all imaging was done at room temperature, and that was measured intermediately during imaging experiments. For the analysis of cell movement and the purpose of reading oscillatory signals, temperature control was achieved using a Bachhoffer chamber and the temperature was maintained at 28°C. The temperature inside the petri-dish during imaging was measured with a K-type thermocouple dipped into the E3 (without methylene blue) medium and the temperatures were recored every 5 minutes.

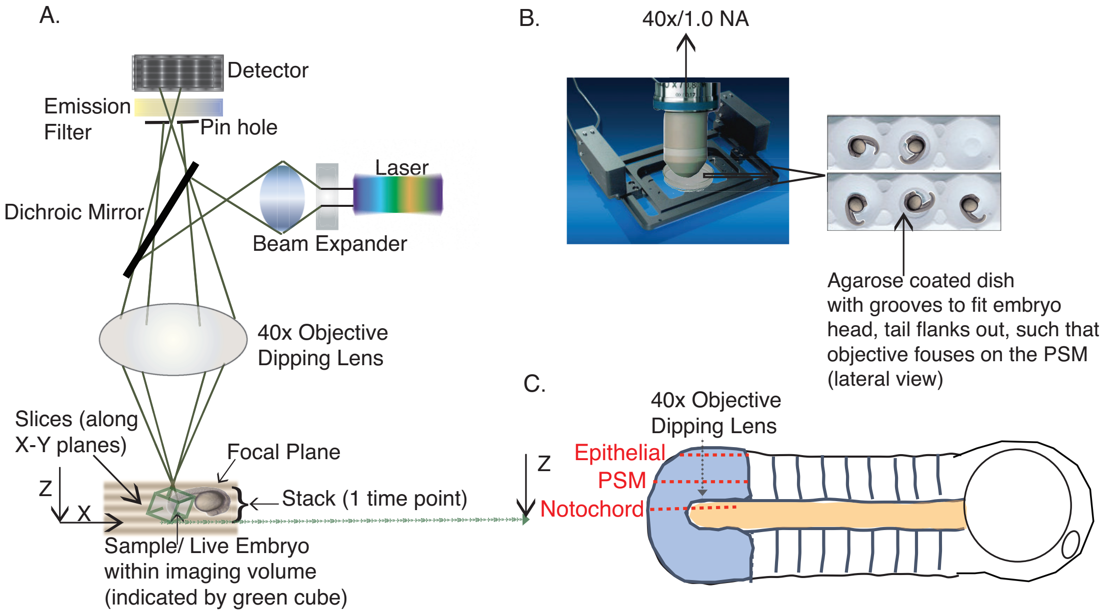


Figure 6.1: A. Illustration of the imaging technique used in this work. The width of the pin hole determines the amount of light that reaches the sample. A 40x Dipping lens is used to image the PSM of the live embryo over several xy planes to obtain a 3D stack. B. Sample setup. A petri-dish coated with agarose with conical depressions such that embryo head fits laterally and the tail flanks out parallel to the agarose. The objective is immersed into the dish focussed on posterior body of the embryo. C. Imaging the sample. The embryo is imaged over time laterally from the epithelial tissue, through the PSM, until the notochord nuclei are visible. The imaging frame is fixed.

6.4 Nuclei Segmentation Algorithm

Images were segmented using custom-written Matlab programs that were exported as a plug-in 'CWNT' or 'Crown wearing nuclei tracking' in Fiji, an open source image analysis software (Schindelin et al., 2012). Fig. 4.4 briefly outlines the nuclei segmentation scheme. We have built a 3D image analysis tool that aims to recognize in vivo juxtaposed nuclei over time. The algorithm is fast and accurate, and is based on combining image derivative followed by carefully inspecting features of 3D segmented objects to enhance the results. The algorithm has several tunable parameters and is designed to work for biological image data typically labeled with nuclear markers.

6.4.1 Input Image De-noising

From the 3D stack, we take the 2D image slice for de-noising. Image de-noising is a crucial step in image analysis. This essentially should remove noisy pixels and at the same time not degrade features of interest in the image. Here, we list a few linear and non-linear filters that can potentially be useful to reduce noise in the images (Lim, 1990; Gonzalez and Woods, 1992). Since our algorithm relies on sharp cell edges, we prefer to de-noise images such that we dilute the noisy component without degrading the change of image gradients at the edges.

Median filter: A non-linear filter, effective for salt and pepper noise and speckled noise. It preserves the edges for a given window size. The value of an output pixel is determined by the median of the neighborhood pixels.

Wiener filter: A pixel-wise adaptive Wiener method based on statistics estimated from a local neighborhood of each pixel. It can be used for removal of blur in images due to power additive noise.

Gaussian filter: A low-pass filter linear smoothing operation that convolves the image with a Gaussian function and brings the value of each pixel into closer harmony with the values of its neighbors (Shapiro and Stockman, 2001).

Bilateral filter: A non-linear, edge preserving and noise reducing smoothing filter. The principle here is a convolution operation with a non-linear Gaussian filter with weights based on pixel intensities. It is applied in the form of two Gaussian filters at a localized pixel neighborhood, one in the spatial domain and the other in the intensity domain.

Lucy Richardson Deconvolution filter: It is an iterative Deconvolution algorithm that attempts to de-blur an image. It assumes that the image has been blurred by a point-spread function and pixel statistics follow Poisson distribution, appropriate for photon noise in image data (Richardson., 1972; Lucy., 1974)

The linear and non-linear filters mentioned above were chosen depending upon the type of noise in the image to obtain an *FIM* image, which had reduced high frequency noise. *IM* is the raw input image and *FIM* is the de-noised image obtained after using a de-noise filter operator, *sf.*

$$FIM = sf(IM), \tag{6.1}$$

We typically used the Gaussian filter given by, $g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$, for very low background noise in images with a value σ of 0.5 and window size of 5×5. For images with high noise, we first used the Gaussian filter followed by Deconvolution and median filters.

6.4.2 Non-linear Isotropic diffusion filter

Originally proposed by Perona and Malik (Perona and Malik., 1990), it evolves the image with a smooth partial differential equation, similar to the heat equation and yields intra-region smoothness and at same time impedes diffusion at the edges. We apply the filter *df* proposed by Perona and Malik on the de-noised image as it facilitates edge extraction, before computing image derivatives. We implemented the standard filter given by,

$$\frac{\partial FIM}{dt} = \mathrm{div}[df(||\nabla FIM||)\nabla FIM], \tag{6.2}$$

where $||\nabla FIM||$ is the de-noised image intensity gradient and *df* is the diffusion co-efficient, given by, $df(||\nabla FIM||) = \exp(-||\nabla FIM/\kappa^2||)$ in this work. This step smoothened noisy pixels, although features of interest in the image required for segmentation (edges of nuclei) were not diluted. I set the value of κ to 50 for about 10 iterations for the purpose of our image analysis.

6.4.3 Computation of Image Derivatives

Image Gradient: first derivative of the image

After image de-noising and implementing the anisotropic diffusion filter, the image edge pixels are enhanced compared to the background and pixels within the nucleus. Hence, I computed the gradient of the image, where the partial derivative, ∂x is obtained by convolving the image with the derivative of the Gaussian function in the x direction and the Gaussian function itself in the y direction and vice-versa for ∂y. Hence, the width of Gaussian parameter allows the tuning of the gradient norm (Shapiro and Stockman, 2001; Canny., 1986).

$$\nabla ADF = [\partial_x ADF \;\; \partial_y ADF], \tag{6.3}$$

For further analysis, I used the magnitude of the gradient given by:

$$||ADF|| = \sqrt{\partial_x ADF^2 + \partial_y ADF^2}, \tag{6.4}$$

Image Laplacian and Hessian: Second derivative of the image

I re-differentiated the result of the gradient of the image and obtain the second order partial derivatives, given by the Hessian matrix.

$$H_m = \begin{pmatrix} \frac{\partial^2 ADF}{\partial x^2} & \frac{\partial^2 ADF}{\partial x \partial y} \\ \frac{\partial^2 ADF}{\partial x \partial y} & \frac{\partial^2 ADF}{\partial y^2} \end{pmatrix}, \tag{6.5}$$

The trace of the matrix is the Laplacian of the image, given by

$$L_a = \frac{\partial_2 ADF}{\partial x^2} + \frac{\partial_2 ADF}{\partial y^2}, \tag{6.6}$$

The Laplacian ingredient for our final filter is a function L,

$$L = \begin{cases} L_a, \text{ if } L_a > 0 \\ 0, \text{ otherwise} \end{cases}, \tag{6.7}$$

Our main objective is to separate closely spaced nuclei. For this, we used the determinant of H_m that marks the saddle points. This component efficiently cuts through the touching spaces between neighboring nuclei, thus contributing to the robustness of segmentation results. The function H marks all the saddle points in the image, given by:

$$H = \begin{cases} |det(H_m)|, \text{ if } det(H_m) < 0 \\ 0, \text{ otherwise} \end{cases}, \tag{6.8}$$

Combining Derivatives

At this point, I combined equations (6.4), (6.7) and (6.8) with different weighing constants α, β, ϵ in order to obtain a function that is large in the form of a ring-shape given by D, where,

$$D = \alpha||\nabla ADF|| + \beta L + \epsilon H, \tag{6.9}$$

In order to obtain a smooth edge for every object, I used a tangent hyperbolic function. The touching points between nuclei form the saddle points of the image intensity matrix, ADF, characterized by $det(H_m) < 0$. Combining derivatives, with this tangent hyperbolic function, I defined a masking function F,

$$F = 0.5(tanh(X) + 1), \tag{6.10}$$

Where ,

$$X = [1 - \frac{D}{\gamma\delta}], \tag{6.11}$$

The parameters $\alpha, \beta, \epsilon, \gamma$ and δ can be tuned to achieve good segmentation results. In the end, I do a mathematical multiplication of the masking function F with the de-noised image *FIM* to obtain a matrix *FI* that has highlighted edges, given by,

$$FI = F.FIM. \tag{6.12}$$

This image is thus prepared for the subsequent image thresholding step.

6.4.4 Image Thresholding

Based on the sequence of steps described so far, I obtain an image *FI*, with clearly worked out the nuclei edges. Further, I used Otsu's thresholding method (Otsu., 1979) that finds suitable threshold value for every 2D slice, each varying in nuclear density, contrast and packing. The Otsu's method finds a suitable threshold for each slice by first representing two classes of pixels, (foreground and background) as a bi-modal histogram and then calculating the optimum threshold separating the two classes so that their intra-class variance is minimal.

6.4.5 From 2D to 3D: Connecting the like pixels

2D binary slices are connected such that the neighboring pixels with value 1 are clustered together to make 3D objects and with values 0 contribute to the background voxels (Haralick and Shapiro, 1992). I used the minimum connectivity of 6 neighboring pixels to make up 3D voxels (background and foreground objects). This connectivity to optimize the total number of segmented objects obtained. I found that for higher order of connectivities consisting of 18 or 26 neighbors, the number of segmented objects decreases, indicating that these lead to severe under-segmentation problems. At this point, I have completed the image segmentation. I now further examine the 3D objects based on the properties of each segmented object.

Analysis of Properties of 3D Segmented nuclei

Each 3D segmented object was characterized by three key features, namely,
Voxel list: The voxel coordinates in x, y and z make up the voxel list for the segmented object.
Centroid: The mean of voxels in each direction gives the centroid of that object.
Volume: The total number of voxels that make up the segmented nuclei.

6.4.6 Post-processing based on properties of 3D segmented nuclei

I used the histogram of the volumes of objects as a tool to draw a boundary between correctly segmented and under-segmented objects. The histogram plot of 3D volumes is positively skewed exhibiting that even after employing a rigorous segmentation algorithm that specifically aimed to separate juxtaposed nuclei, it is possible to have multiple fused objects. The imaging volume spans over different tissues of the developing embryo. Therefore a wide range of heterogeneous shapes and sizes of nuclear volumes is observed even for correctly segmented objects, as depicted in the histogram. In the case of under segmentation, there would be particularly large volumes, that are in fact comprised of several nuclei with only a single centroid. I empirically determined a cut-off representing the maximum allowed volume for correctly segmented objects, given by,

$$FOC = MV + (0.1 \times SDV) \tag{6.13}$$

where FOC is the fused object Cut-off, MV is the mean volume of segmented objects and SDV is the standard deviation of volumes of segmented objects.

For all potentially fused objects, I sampled the voxel frequency distribution in each x, y and z direction separately. Ideally, one would expect to see a uni-modal distribution with one local maximum in each list (for a typical image of a nucleus obtained from the LSM) and the geometrical centroid after segmentation should lie at the centre of nuclear volume as shown in Fig. 4.8 . In contrast, fused objects would have multi-modal frequency distributions with multiple local maximas. This information can now be further used to split fused objects into potential nuclei candidates. For this, I devised a semi-automated method to re-segment the fused objects by using the local peak information from the coordinate voxel distributions.

Based on the total number of local maxima in the frequency distribution of voxels in the three directions, x, y and z, we formed a guess for the number of potentially fused nuclei clustered together. This value was provided to the k-means that was able to segregate the voxels into clusters with their corresponding new centroids and new voxel list values.

Each cluster in this case, represents a true nucleus. K-means clustering uses the Euclidian distance metric and the variance between the data points and determines cluster centroids (Seber, 1984; Spath, 1985; Hartigan and Wong, 1979; Dhillon and Modha, 2001; Amorim and Mirkin, 2012; Press et al., 2007). An important feature of k-means is that it requires the actual number of clusters as an input parameters, where an inappropriate choice leads to the wrong number of clusters, since in our case, this parameter is based local maxima from the voxel list. It is vital to reduce noise from the image, while keeping smooth intra-nuclear regions, to avoid false local maximas.

Alternatively, I propose the use of Gaussian mixture models (GMM) (McLachlan and Peel, 2000; Bishop, 2006; Press et al., 2007) for post-processing nuclei clusters. GMM were also initialized with the number of local maxima as a guess for the number of nuclei. This value was used as the number of components required to fit parametric models to the data. Choosing a suitable number of components or nuclei clusters is essential for creating a useful model of the data (voxel list) - too few components fails to model the voxel list accurately; too many components leads to an over-fit model with singular covariance matrices. Therefore, I combined GMM with Akaike information criterion (AIC) (Akaike, 1980; Figueiredo and Jain, 2002), which penalizes the model based on its complexity. The GMM fits several models with different proportions to the same voxel list and subsequently AIC minimizes the number of 3D Gaussian models and ultimately automatically achieves the correct number of Gaussian models or nuclei, even in cases where the initial guess was slightly off.

6.5 Generation of Synthetic Images

Synthetic images were used to assess the systematic errors in the algorithm in a case where we exactly know the centroids of objects (Sbalzarini and Koumoutsakos, 2005). Synthetic images were generated by Dr. Koichiro Uriu, a post-doc from our group.

We randomly allocated centroids of N objects in a three-dimensional continuous space $L_x \times L_y \times L_z$ (in µm). Let $x_j = (x_j, y_j, z_j)$ be the centroid of an object j ($j = 1, 2, .., N$) in the three-dimensional space. We defined density of objects

as ρ=N/ $(L_x\,L_y\,L_z)$. We chose the number of objects N such that the density ρ in synthetic images is close to that of actual PSM images. In order to obtain a preferred distance between objects, we relaxed their locations by using equation of motions with a Lennard-Jones potential (Frenkel and Smit, 2002).
We approximated the shape of nuclei by an ellipsoid with shape parameters a, b, and c, which also determine their volume. We set a = 4µm, b = 2.5µm, and c = 2.5µm based on the measured average size of a typical nucleus from the embryonic data. The orientation of the object j is given by the Euler angle $(\alpha_j, \beta_j, \gamma_j)$. This angle was chosen from a uniform distribution between 0 and $\pi/4$ for α and β and between 0 and $\pi/2$ for γ.

To represent the orientation of the object j, we rotated its axes such that:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_j & -\sin\alpha_j \\ 0 & \sin\alpha_j & \cos\alpha_j \end{bmatrix} \begin{bmatrix} \cos\beta_j & 0 & \sin\beta_j \\ 0 & 1 & 0 \\ -sin\beta_j & 0 & \cos\beta_j \end{bmatrix} \begin{bmatrix} \cos\gamma_j & -\sin\gamma_j & 0 \\ \sin\gamma_j & \cos\gamma_j & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x - x_j \\ y - y_j \\ z - zj \end{bmatrix} \tag{6.14}$$

The signal intensity at a given position (x, y, z), emitted by the object j, $I_j(x, y, z)$ was modeled as follows:

$$I_j(x, y, z) = \frac{I_0}{2I_n}(\tanh[\sigma 1 - \frac{x'^2}{a^2} - \frac{y'^2}{b^2} - \frac{z'^2}{c^2}] + 1), \tag{6.15}$$

where I_0 is the maximum intensity at the centroid of the object, s represents the steepness of the edge of an object and I_n is the normalization constant, $I_n = (\tanh(\sigma) + 1)/2$. The intensity value at the position *(x, y, z)* is, given by $I(x,y,z) = I_b + \max[I_j(x, y, z), |j = 1, 2, N]$, where I_b is the background intensity. To determine I_0 and I_b, we measured the average intensity at the center of nucleus I_m and background intensity $\hat{I}_b$ in our real images. Next, we set $I_0 = \hat{I}_b$ and $I_0 = I_m - \hat{I}_b$.
The three-dimensional space was discretized using the voxel spacing in each direction and images were created from the intensity profiles. Let D_x, D_y and D_z be the voxel sizes in x, y and z directions, respectively. The voxel size represents the

spatial resolution of an image. We set the same voxel size as obtained from our confocal microscope images ($D_x = 0.691\mu m, D_y = 0.691\mu m, D_z = 1.75\mu m$). The discrete spatial coordinate of the three-dimensional space $(\tilde{x}, \tilde{y}, \tilde{z})$ can be written as $(\tilde{x}, \tilde{y}, \tilde{z}) = (k_x\Delta x,\ k_y\Delta y,\ k_z\Delta z)$ where, $k_\zeta = 1, 2, , L_\zeta/\Delta_\zeta$ and $x = (x,\ y\, or\, z)$. The signal intensity of the voxel (k_x, k_y, k_z) in the synthetic image was given by:

$$I_d(k_x, k_y, k_z) = I(\frac{\Delta x}{2}(2k_x - 1),\ \frac{\Delta y}{2}(2k_y - 1),\ \frac{\Delta z}{2}(2k_z - 1)) \tag{6.16}$$

We then added the noise to each voxel in the synthetic image, we used gamma distribution, given by:

$$f(x, k, \theta) = x^{k-1}\frac{e^{-x/\theta}}{\Gamma(k)\theta^k}, \tag{6.17}$$

where, k and θ are the shape and scale parameters, respectively and Γ represents the gamma function. The mean of the gamma distribution is $m = k\theta$ for which we choose the intensity at each voxel, $k\theta = I_d(k_x, k_y, k_z)$. We defined the SNR as the inverse of the coefficient of variation (CV) given by: $SNR = 1/CV = m/\sigma = \sqrt{k}$. The shape and scale parameters, may be written in terms of a given SNR as $k = SNR^2$ and therefore, $\theta = I_d(k_x, k_y, k_z)/SNR^2$. Note that q might differ between two different voxels under this setting depending on their noiseless intensities $I_d = (k_x, k_y, k_z)$. We generated a random real number from the gamma distribution with k and q obtained by the above equations for each voxel. Finally, 16-bit images were produced with the 'Image' function in Mathematica from the intensity profile.

6.5.1 Segmentation efficacy measured by Sensitivity and Precision

We created artificial datasets by varying the SNR and density to evaluate the algorithm performance. To measure error rates in our segmentation algorithm, we matched each of the data points obtained by the segmentation algorithm with each of the true positions based on Euclidian distances between them. We adopted particle-matching algorithm previously used for cell tracking and minimized the cost function defined as the summation of Euclidian distances between assigned pairs (Sbalzarini and Koumoutsakos, 2005).

We compared segmentation results of the synthetic data with their known true centroid positions. If a match was found between true position of an object and the segmented position, the algorithm was qualified to have segmented the object correctly. A miss was considered when a true position was not paired with any of the segmented data points. On the other hand, when a segmented data point was not paired with any true positions, we considered that the segmentation algorithm detected false signals (e.g. noise) as an object or over-segmented an object. The mis-matches and the false signals were accounted in the form of false positives (over segmentation rate) and false negatives (under segmentation rate). In order to measure the error rates by the algorithm, we defined sensitivity and precision.

We counted the number of real objects correctly detected by the algorithm N_{ra} and defined sensitivity as N_{ra}/N_r where N_r is the true number of objects in a synthetic image. The sensitivity represented the detection rate of objects in an image. We defined precision as N_{ra}/N_s where N_s is the total number of segmented objects. The precision represents the probability that a detected object is actually a real object. The score was given by the product of sensitivity and precision.

6.6 Transplant Experiments

Genetic mosaics were generated as described previously by (Westerfield, 2007; Haas and Gilmour, 2006). Briefly, donor embryos expressing two histone variants fused to gfp and m-cherry (h2AflV-gfp, h2AflV-mcherry) were allowed to develop until blastula period. Approximately 20 to 30 cells were then transplanted into age-matched h2AflV-gfp host embryos. The next day, mosaic h2AflV-gfp embryos were screened for mcherry expression and imaged with a Zeiss LSM 780 equipped with a 40x dipping objective/N.A.1.0. Fig. 6.2 shows a brief sketch of the transplantation experiments done to obtain chimeric embryos used here to examine segmentation algorithm results in live embryo.

6.6.1 Segmentation performance assessment using embryo chimeras

We selected small cubes in different regions of the embryo and segmented nuclei in both dense and sparse channels Fig. 4.13B. The sparsely labeled nuclei channel

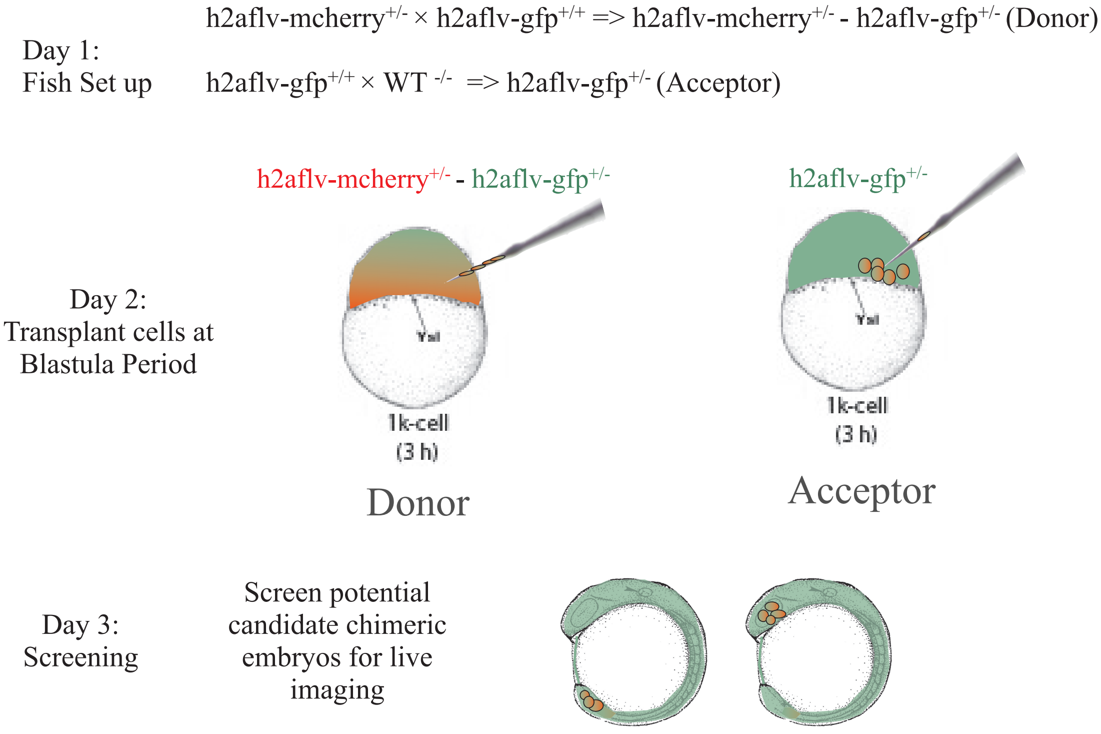


Figure 6.2: Transplant experiment procedure. Fish crosses were done to obtain 'donor embryos' with double labeled ubiquitous nuclei markers- mcherry and gfp and 'acceptor Embryos' with single labeled gfp nuclei marker. Cells from the donor embryos were transplanted into acceptor embryos around blastula period. Embryos were screened on the next day for live imaging

allowed us to achieve reliable segmentation results. We computed the Euclidian distance between all the centroids from the two channels. A match between two nuclei from each channel was considered when the distance was below 2µm. We defined the match fraction as the ratio of the ‘number of matches between dense and sparse‘ to the ‘number of nuclei in the sparse channel‘.

6.7 Nuclei Tracking Algorithm

I adopted the trajectory linking algorithm previously written for particle tracking (Sbalzarini and Koumoutsakos, 2005). The algorithm uses a association matrix based on Euclidean distance between nuclei of consecutive time frames. A cost function minimizes the distances considering all the time frames.The nearest neighbor linking distance was tested for a range of values between 3.5-5µm and a value of 5 µm was chosen for linking trajectories. Further, embryo chimeras were analyzed in order to understand the relationship between density of nuclei and trajectory lengths.

6.8 Measurement of Diffusion co-efficient of cells

A method was developed in this work to measure the diffusion coefficient of cells in different regions of the PSM of the embryos along the anterior-posterior axis, on the lateral line. 3D segmented nuclei centroids from the first time point were rendered and a single centroid was selected manually. For the selected centroid position, we select all the tracks originating from the first time point that lie within a radius of 18µm. Fig. 6.3 illustrates the region selection procedure explained so far for the analysis.

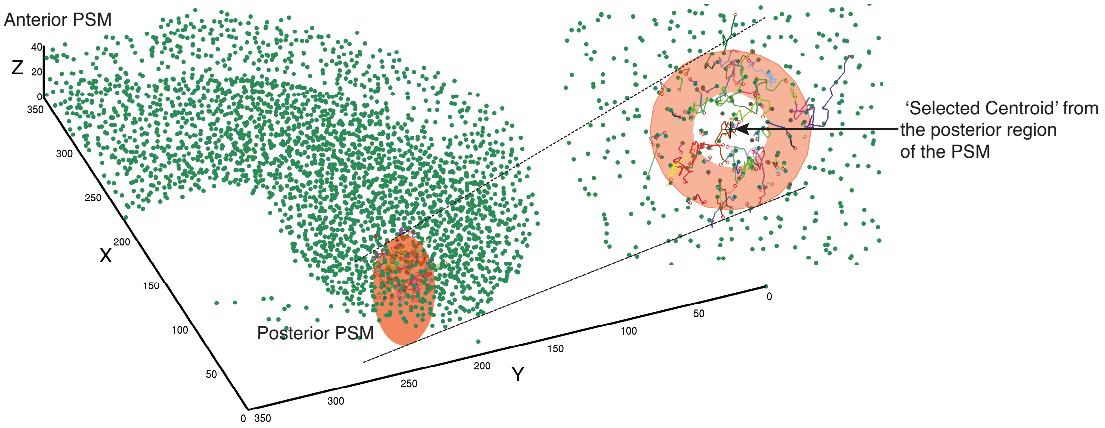


Figure 6.3: 3D cell tracks were extracted around the selected centroid, indicated by a black asterisk marker, within a 2d radius of 18µm' as highlighted by the orange ellipsoid in the posterior PSM, for the purpose of analysis of diffusion coefficient of cells. Snapshot from a 19 somite stage WT embryo imaged at 28°C with time interval between consecutive frames of 1.2 minutes. On the right, a zoomed in, 2D projection of the selected region is shown.

Since the selected centroid tracks varied in length, I processed the samples to obtain one with correct number of members, such that all have sufficient length and provide enough statistics.

For this, I defined a fractional cut-off, f_c of the original nuclei that are supposed to remain in the sample. Large f_c thus ensures good statistics. We then determined the time sample t_n where, $N(t_n) = f_c.N_0$, where N(t) is the time varying number of nuclei in our initial selection and $N(t_1) = N_0$. We then selected only those tracks that extend until t_n. By doing this, we avoided large standard error of mean in the calculation of Mean square displacement owing to lesser number of tracks at later time steps. Therefore, a selected set of spatial coordinates of nuclei was examined to determine a useful f_c and $t_1, t_2, ...t_n$ from a selected region of the tissue was used for further analysis.

For the calculation of MSD, each track was expressed relative to its first time sample, $(\tilde{x}_i(t_n),\, \tilde{y}_i(t_n),\, \tilde{z}_i(t_n)) = (x_{tn} - x_{t1}), (y_{tn} - y_{t1}), (z_{tn} - z_{t1})$, where *t1* stands for first time sample and *tn* for time samples from 2 to n (last time sample), for a cell, i.

Pair-wise distances were computed at each time step given by,

$d_{ij}^2 = (\tilde{x}_i(t_n) - \tilde{x}_j(t_n))^2 + (\tilde{y}_i(t_n) - \tilde{y}_j(t_n))^2 + (\tilde{z}_i(t_n) - \tilde{z}_j(t_n))^2$, where *i, j* indicate $cell_i$ and $cell_j$.

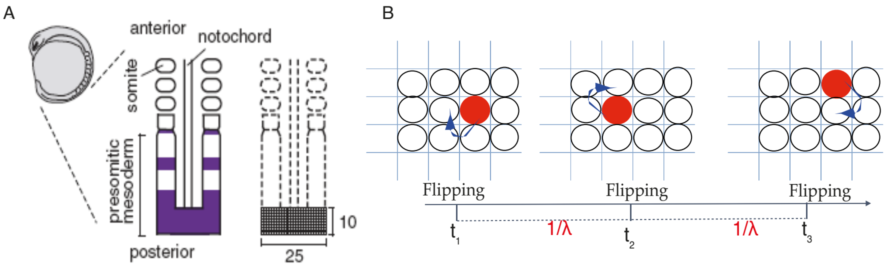
By taking pairs of distances, we essentially eliminated tissue level movements and considered contribution of only cell movements (Gerlich and Ellenberg., 2003).

For each time step, averaged squared distances for all nuclei was computed, given by, $\bar{d_t^2} =< d_{ij}^2 >$.

$\bar{d_t^2}$ gives the Mean squared distance (MSD) as a function of time for the selected nuclei, from a selected spatial region within the PSM tissue.

In order to infer the properties of the PSM cellular oscillators and compute their moving rates, we considered the theoretical model proposed in (Uriu et al., 2010) as shown in Fig. 6.4, that allows cellular oscillators to exchange their positions with their neighbors, which thereby considerably enhances their synchronization properties.

With the model outlined in Fig. 6.4, the time taken for a cell to exchange its location with its neighbor, T_s, encapsulates the time scales of mobility of the cellular oscillators within a specified spatial region of the PSM of the embryo.



Model incoorporates:

Oscillating expressions of Her protein and her gene as a result of intra-cellular negative feedback loop.

Concentration of delta proteins between 4 adjacent neighbors as a result of cellular interactions due to Delta-Notch coupling.

Exchange of cell poistions with their neighbors that leads to decrease in synchronisation time of cellular oscillators than otherwise.

Figure 6.4: A. 2D lattice model for the Posterior PSM of the zebrafish embryo during somitogenesis. Taken from (Uriu et al., 2010). B. The model incorporates features of the *in vivo* events, as indicated in the figure. Cells sitting on the lattice flip their positions with their neighbors between two time points indicated by blue arrows. $1/\lambda$ indicates average time taken for a cell to exchange its position with its neighbor between two time points.

$$T_s = \frac{1}{\lambda} \tag{6.18}$$

Our observation indicates that the diffusion of cells along the anterior-posterior axis seemingly varies within the PSM, with high mixing in the posterior and relatively less in the anterior PSM. Combining the know-how regarding the mobility of oscillators and the model properties, we were interested to explore the moments of displacements that has been previously shown to provide a complete characterization of diffusive behaviors in different biological and physical systems (Ferrari et al., 2001).

Therefore, we considered the PSM cells to be diffusing according to the function,

$$MSD = Dt^{\alpha}. \tag{6.19}$$

For a cell to exchange its position with its neighbor, it has to diffuse about one cell diameter, $C_d = 10\mu m$. Therefore, the switching time, t_s, taken for a cell to exchange its position with its neighbor, can be derived by considering,

$$MSD = C_d^2. \tag{6.20}$$

Combining (6.19) and (6.20), we search,

$$C_d^2 = Dt_s^{\alpha}. \tag{6.21}$$

From here, it was possible to extract, t_s, given by,

$$t_s = (\frac{C_d^2}{D})^{\frac{1}{\alpha}} \tag{6.22}$$

$$\lambda = \frac{1}{2t_s}, \tag{6.23}$$

where, $T_s = 1/\lambda$.

Thus, we can extract the average time, T_s taken for a cell to exchange its position with its neighbor. A Matlab power law fit function was used, on the computed MSD data in order to derive the values of D and α.

Using this method, it was possible to extract diffusion coefficient of cells and their respective time scales in different regions of the PSM along the anterior-posterior axis of the WT as well as mutant embryos. Further, it was also possible to extract the value of α along the anterior-posterior axis of the embryo, that may provide a vivid picture on how cell diffusivity changes along the embryonic axis, and its sensitivity towards temperature and the developmental stage of the PSM. This further allowed us to use this method of analysis of cell movements in mutant embryos, such as *Trilobite*.

The power law fits were tested for a range of values of f_c and fits were evaluated using goodness-of-fit Statistics in Matlab. The final values of D, α and T_s that fit with minimum error were chosen, based on the following measures:

Sum of Squares Due to Error (SSE): This measures the total deviation of the data values from the fit. A value closer to 0 would be indicative of a model with smaller error component.

$SSE = \sum_{i=1}^{n} w_i(y_i - \hat{y_i})^2$, where w_i are the weights, y_i is the data and $\hat{y_i}$ is the fit.

Degrees of Freedom Adjusted R-Square: The adjusted R-square statistic is an excellent indicator of the fit quality for comparing a series of models.

$AdjustedR^2 = 1 - \frac{SSE(n-1)}{SST(v)}$

where SST gives the sum of squares about the mean, $SST = \sum_{i=1}^{n} w_i(y_i - \bar{y_i})^2$. The residual degrees of freedom (v) is given by the number of response values (n) minus the number of fitted coefficients (m) estimated from the response values. The adjusted R-square statistic can take on any value less than or equal to 1, with a value closer to 1 indicating a better fit. Negative values can occur when the model contains terms that do not help to predict the response.

Root Mean Squared Error: This statistic is also known as the fit standard error and the standard error of the regression. It is an estimate of the standard deviation of the random component in the data, and is defined as

$RMSE = \sqrt{(MSE)}$, where $MSE = \frac{SSE}{v}$, and MSE is the mean square error or the residual mean square. Just as with SSE, an MSE value closer to 0 indicates a fit that is more useful for prediction. For more details on this section, refer to Matlab Help on Goodness-of-fit Statistics.

The method was further modified to verify diffusion of cells along the lateral line. This was essentially done by selecting 3D sections of the PSM tissue and computing diffusion within 3D radius of 18µm around each selected cell within the PSM space. This allowed us to inspect time scales of diffusion for the entire image along the anterior-posterior axis of the embryo as well along the dorso-ventral direction on one of the lateral sides of the embryo.

6.9 Reading oscillating expression using nuclear marker

Embryos, carriers of *her1*-yfp and nuclear marker, h2aflv-mcherry were imaged by sequential scanning and further the cherry channel nuclei were segmented using the CWNT algorithm and nuclei traces were made using the tracking algorithm. The centroids as well as the voxel list values for each tracked nucleus were registered for the mcherry channel. The voxel list locations were used to read the oscillating *her1* expression of the yfp channel. This method essentially allowed us to read the time series trajectory of an oscillating cell using the mcherry nuclear marker. The average fluorescence intensity (i.e. the average voxel intensity) within each nucleus was computed as the total nuclear fluorescence intensity divided by the volume of the nucleus and plotted as a function of time to visualize the dynamics of the oscillations.

Chapter 7

Supplementary Movies

Movie S1: Lateral view of a 18 somite stage WT zebrafish embryo, taken from a Zeiss Meta 510, Confocal upright microscope, using a 40x Dipping lens, 1.0 NA. A stack of 8-bit image ($512 \times 512 \times 39$) acquired using 488 laser that excites H2aflv-gfp line; 1 voxel given by : $0.621 \times 0.621 \times 1.745$ µm^3, image acquired at time interval of 81.6 seconds.

Movie S2-A: Lateral view movie of the 18 somite stage embryo (in S1) at 6.98 µm in z-plane (4th slice movie at 7 frames per second, using Fiji)

Movie S2-B: Lateral view movie of the 18 somite stage embryo (in S1) at 34.9 µm in z-plane (20th slice movie at 7 frames per second, using Fiji)

Movie S2-C: Lateral view movie of the 18 somite stage embryo (in S1) at 68.05 µm in z-plane (39th slice movie at 7 frames per second, using Fiji)

Movie S3-A: 3D movie (using 3D Fiji viewer) of 19 somite stage WT zebrafish embryo, taken on a Zeiss Zen 710, single photon upright microscope, using a 40x Dipping lens, 1.0 NA, GaAsP spectral detector. A stack of 16-bit image ($512 \times 512 \times 23$) acquired using 488 laser that excites h2aflv-gfp; 1 voxel given by : $0.621 \times 0.621 \times 1.79$ µm^3, image acquired at time interval of 75.3 seconds.

Movie S3-B: 3D movie of the 19 somite stage zebrafish embryo (in S3-A) after image segmentation using the proposed algorithm. Movie shows rendered nuclei in green after image segmentation over 70 time frames (using Matlab).

Movie S4: 3D movie (using 3D Fiji viewer) of a 18 somite stage WT zebrafish embryo, taken on a Zeiss Meta 510, Confocal upright microscope, using a 40x Dipping lens, 1.0 NA. A stack of 8-bit image ($512 \times 512 \times 39$) acquired using 488 laser that excites h2aflv-gfp line; 1 voxel given by : $0.621 \times 0.621 \times 1.75$ μm^3 with time interval between frames being 130 seconds. Embryo outlined in green color in the first frame and by the end of the movie, the embryo rolls over during imaging because of its growth. Embryo is outlined by a red color line in the last frame of the movie. Image data was used for verification of total number of segmented nuclei in comparison with results from Imaris software.

Movie S5: Transplant experiment– 3D movie (using 3D Fiji viewer) of a 16 somite stage WT embryo chimera after transplantation experiment, taken on a Zeiss 710 Zen, single photon upright microscope, using a 40x Dipping lens, 1.0 NA, GaAsP spectral detector. A stack of 16-bit image ($512 \times 512 \times 35$) was acquired using 488 and 561 lasers that excite h2aflv-gfp and h2aflv-mcherry respectively; 1 voxel given by : $0.69 \times 0.69 \times 1.75$ μm^3, image acquired with time interval between frames being 72.6 seconds.

Movie S6: Transplant experiment– 3D movie (using 3D Fiji viewer) of a 17 somite stage WT embryo chimera after transplantation experiment, taken on a Zeiss 710 Zen, single photon upright microscope, using a 40x Dipping lens, 1.0 NA, GaAsP spectral detector. A stack of 16-bit image ($512 \times 512 \times 30$) was acquired using 488 and 561 lasers that excite h2aflv-gfp and h2aflv-mcherry respectively; 1 voxel given by : $0.69 \times 0.69 \times 1.90$ μm^3,image acquired with a time interval between frames being 65.7 seconds.

Movie S7-A: 3D movie (using 3D Fiji viewer) of a 19 somite stage *Trilobite* mutant zebrafish embryo imaged at 28°C, taken on a Zeiss 710 Zen, single photon upright microscope, using a 40x Dipping lens, 1.0 NA. A stack of 16-bit image ($512 \times 512 \times 38$) acquired using 488 laser that excites h2aflv-gfp; 1 voxel given by

: $0.691 \times 0.691 \times 1.79$ µm^3, image acquired at time interval of 79.1 seconds with 70 time steps.

Movie S7-B: 3D movie (using 3D Fiji viewer) of a 17 somite stage *aei* mutant. Microinjections with mkate2-nls nuclear marker around 8-cell stage was done and the mutant zebrafish embryo was imaged at 22°C, taken on a Zeiss 710 Zen, single photon upright microscope, using a 40x Dipping lens, 1.0 NA. A stack of 16-bit image ($512 \times 512 \times 25$) acquired using 561 laser; 1 voxel given by : $0.691 \times 0.691 \times 1.80$ µm^3, image acquired at time interval of 42.5 seconds with 70 time steps.

Movie S8-A: WT *her1* 18 somite stage embryo, imaged at 22 °C, taken on a Zeiss Meta 510, Confocal upright microscope, using a 40x Dipping lens, 1.0 NA. A stack of 8-bit image ($512 \times 512 \times 20$) acquired using 514 laser that excites yfp; 1 voxel given by : $0.62 \times 0.62 \times 2.10$ µm^3 at time interval of 160.5 seconds with 40 time steps. Movie shows waves of oscillating *her1* expression traversing the tissue from the posterior PSM to the anterior PSM on a 2D slice at 21 µm.

Movie S8-B: WT embryo at 25 somite stage, labeled with h2aflv-mcherry marker that ubiquitously marks all the nuclei and *her1* labeled with de-stabilized yfp marker. Embryo imaged at 28 °C on a Zeiss 710 Zen, single photon upright microscope, using a 40x Dipping lens, 1.0 NA. A stack of 16-bit image ($512 \times 512 \times 19$) acquired using 561 and 514 lasers that excites mcherry and yfp respectively. 1 voxel given by : $0.69 \times 0.69 \times 2.0$ µm^3 at time interval of 89 seconds with 20 time steps.

Movie S9-A: *her1*-yfp oscillations on a *Trilobite* background at 17 somite stage embryo imaged at 28 °C, on a Zeiss 710 Zen, single photon upright microscope, using a 40x Dipping lens, 1.0 NA. A stack of 16-bit image ($512 \times 512 \times 34$) acquired using 514 laser that excites yfp; 1 voxel given by : $0.62 \times 0.62 \times 2.0$ µm^3 at time interval of 71 seconds with 80 time steps. Traveling waves sweep across the tissue from posterior to anterior. Note the inter-stripe distance is smaller compared to WT as well as the size and length of the *her1* stripe.

Movie S9-B: *Trilobite* embryo at 17 somite stage, labeled with h2aflv-mcherry marker that ubiquitously marks all the nuclei and *her1* labeled with de-stabilized yfp marker. Embryo imaged at 28 °C on a Zeiss 710 Zen, single photon upright microscope, using a 40x Dipping lens, 1.0 NA. A stack of 16-bit image ($512 \times 512 \times 31$) acquired using 561 and 514 lasers that excites mcherry and yfp respectively. 1 voxel given by : $0.62 \times 0.62 \times 2.0$ μm^3 at time interval of 65.0 seconds with 80 time steps.

Movie S10: A stack obtained from SPIM for an embryo from 18 somite stage onwards, carrier of h2aflv-gfp. The 16-bit stack dimensions are ($960 \times 960 \times 68$), where 1 voxel is given by : $0.279 \times 0.279 \times 0.186$ μm^3. Image is acquired using a 25x lens at a frame rate of 30.5 seconds using a 488 laser line.

Bibliography

Akaike H; *Bayesian Statistics: Likelihood and the Bayes procedure* (Valencia: University Press,, u, 1980); 1st edition.

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, and Walter P; *Molecular Biology of the Cell* (New York: Garland Science, USA, 2002); 4th edition.

Amorim RC and Mirkin B; *Minkowski metric, feature weighting and anomalous cluster initializing in k-means clustering.*; Pattern Recognition; **45**, 1061; 2012.

Amos WB and White JG; *How the confocal laser scanning microscope entered biological research.*; Biology of the Cell; **95**, 335; 2003.

Ashe HL and Briscoe J; *The interpretation of morphogen gradients.*; Development; **133**, 385; 2006.

Aulehla A, Wiegraebe W, Baubet V, Wahl MB, Deng C, Taketo M, Lewandoski M, and Pourquié O; *A beta-catenin gradient links the clock and wavefront systems in mouse embryo segmentation.*; Nat Cell Biol; **10**, 186; 2008.

Baker RE and Maini PK; *Traveling gradients in interacting morphogen systems.*; Math Biosci; **209**, 30; 2007.

Bénazéraf B, Francois P, Baker RE, Denans N, Little CD, and Pourquié O; *A random cell motility gradient downstream of fgf controls elongation of an amniote embryo*; Nature; **466**, 248; 2010.

Bessho Y, Hirata H, Masamizu Y, and Kageyama R; *Periodic repression by the bhlh factor hes7 is an essential mechanism for the somite segmentation clock.*; Genes Dev.; **17**, 1451; 2003.

Bishop C; *Pattern recognition and machine learning.* (New York: Springer,, USA, 2006); 1st edition.

Blanchard GB, Kabla AJ, Schultz NL, Butler LC, B BS, Gorfinkiel N, Mahadevan L, and Adams RJ; *Tissue tectonics: morphogenetic strain rates, cell shape change and intercalation.*; Nat Methods.; **6**, 458; 2009.

Bleau A and Leon LJ; *Watershed-basedsegmentation and regionmerging.*; Computer Vision and Image Understanding; **77**, 317; 2000.

Bollenbach T, Kruse K, Pantazis P, González-Gaitán M, and Jülicher F; *Robust formation of morphogen gradients.*; Physics Review Letters; **94**, 018103; 2005.

Brakenhoff G, Vischerr K, and edited by J Pawley; *Handbook of Biological Confocal Microscopy* (New York: Plenum, USA, 1995); 1st edition.

Canny J; *Computational approach to edge detection.*; IEEE Transactions on Pattern Analysis and Machine Intelligence; **8**, 679; 1986.

Chan and Vese LA; *Active contours without edges.*; IEEE Transactions on Image Processing; **10**, 266; 2001.

Chudakov DM, Matz MV, Lukyanov S, and Lukyanov KA; *Fluorescent proteins and their applications in imaging living cells and tissues.*; Physiol Rev; **90**, 1103; 2010.

Cinquin O; *Repressor dimerization in the zebrafish somitogenesis clock.*; PLoS Comput Biol; **13**, 32; 2007.

Collier JR, McInerney D, Schnell S, Maini PK, Gavaghan DJ, Houston P, and Stern CD; *A cell cycle model for somitogenesis: mathematical formulation and numerical solution.*; J Theor Biol; **207**, 305; 2000.

Cooke J; *Control of somite number during morphogenesis of a vertebrate, xenopus laevis.*; Nature; **254**, 196; 1975.

Cooke J and Zeeman EC; *A clock and wavefront model for control of the number of repeated structures during animal morphogenesis.*; J Theor Biol; **58**, 455; 1976.

Davis RL and Turner D; *Vertebrate hairy and enhancer of split related proteins: transcriptional repressors regulating cellular differentiation and embryonic patterning.*; Oncogene; **20**, 8342; 2001.

del Corral RD, Olivera-Martinez I, Goriely A, Gale E, Maden M, and Storey K; *Opposing fgf and retinoid pathways control ventral neural pattern neuronal differentiation and segmentation during body axis extension.*; Neuron; **40**, 65; 2003.

Delfini MC, Dubrulle J, Malapert P, Chal J, and Pourquié O; *Control of the segmentation process by graded mapk/erk activation in the chick embryo.*; Proc Natl Acad Sci U S A; **102**, 11343; 2005.

Dequéant ML, Glynn E, Gaudenz K, Wahl M, Chen J, Mushegian A, and Pourquié O; *A complex oscillating network of signaling genes underlies the mouse segmentation clock.*; Science; **314**, 1595; 2006.

Dequéant ML and Pourquié O; *Segmental patterning of the vertebrate embryonic axis.*; Nat Rev Genet.; **9**, 370; 2008.

Dhillon IS and Modha DM; *Concept decompositions for large sparse text data using clustering.*; Machine Learning; **42**, 143; 2001.

Dubrulle J, McGrew MJ, and Pourquié O; *Fgf signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal hox gene activation.*; Cell; **219**, 106; 2001.

Dubrulle J and Pourquié O; *fgf8 mrna decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo.*; Nature; **427**, 419; 2004.

Eils R and Athale C; *Computational imaging in cell biology.*; Journal of Cell Biology; **161**, 477; 2003.

Ferrari R, Manfroi AJ, and Young WR; *Strongly and weakly self-similar diffusion.*; Physica D; **154**, 111; 2001.

Figueiredo MAT and Jain AK; *Unsupervised learning of finite mixture models.*; IEEE Transactions on Pattern Analysis and Machine Intelligence; **24**, 381; 2002.

Frenkel D and Smit B; *Understanding Molecular Simulation.* (San Diego: Academic Press,, USA, 2002); 2nd edition.

Gerlich D and Ellenberg J; *4d imaging to assay complex dynamics in live specimens.*; Nat Cell Biol; **5**, 14; 2003.

Gilbert SF; *Developmental Biology* (Sinauer Associates, USA, Sinauer Associates, 2006); 8th edition.

Giudicelli F, Ozbudak EM, Wright GJ, and Lewis J; *Setting the tempo in development: An investigation of the zebrafish somite clock mechanism.*; PLoS Biology; **5**, 1309; 2007.

Goldbeter A; *Biochemical Oscillations and Cellular Rhythms.* (Cambridge University Press,, United Kingdom, 1996); 1st edition.

Gomez C, Ozbudak EM, Wunderlich J, Baumann D, Lewis J, and Pourquié O; *Control of segment number in vertebrate embryos.*; Nature; **454**, 335; 2008.

Gonzalez RC and Woods RE; *Digital Image Processing* (Addison-Wesley Publishing Company, Inc.,, USA, 1992); 1st edition.

Goodwin BC; *Temporal Organization in Cells: A dynamic theory of cellular control processes* (Academic Press, London, 1963); 1st edition.

Gritsman K, Zhang J, Cheng S, Heckscher E, Talbot WS, and Schier AF; *The egf-cfc protein one-eyed pinhead is essential for nodal signaling.*; Cell; **97**, 121; 1999.

Gurdon JB and Bourillot PY; *Morphogen gradient interpretation.*; Nature; **413**, 797; 2001.

Haas P and Gilmour D; *Chemokine signaling mediates self-organizing tissue migration in the zebrafish lateral line.*; Dev Cell; **10**, 673; 2006.

Hans-Hermann GB and Kaether C; *Tissue tectonics: morphogenetic strain rates, cell shape change and intercalation.*; FEBS Letters; **389**, 44; 1996.

Haralick RM and Shapiro LG; *Computer and Robot Vision* (Addison-Wesley,, USA, 1992); 1st edition.

Hartigan JA and Wong MA; *Algorithm as 136: A k-means clustering algorithm.*; Journal of the Royal Statistical Society, Series C (Applied Statistics); **28**, 100; 1979.

Henry CA, Hall LA, Hille MB, Solnica-Krezel L, and Cooper M; *Somites in zebrafish doubly mutant for knypek and trilobite form without internal mesenchymal cells or compaction.*; Curr Biol.; **10**, 1063; 2000.

Henry CA, Urban MK, Dill KK, Merlie JP, Page MF, Kimmel CB, and Amacher SL; *Two linked hairy/enhancer of split-related zebrafish genes, her1 and her7, function together to refine alternating somite boundaries.*; Development; **129**, 3693; 2002.

Herrgen L; *The regulation of segmentation clock period in zebrafish.*; Ph.D. thesis; Technischen Univesität Dresden; 2008.

Herrgen L, Ares S, Morelli L, Schröter C, Jülicher F, and Oates A; *Intercellular coupling regulates the period of the segmentation clock.*; Curr Biol.; **20**, 1244; 2010.

Herrgen L, Schröter C, Bajard L, and Oates AC; *Multiple embryo time-lapse imaging of zebrafish development.*; Methods Mol Biol.; **546**, 243; 2009.

Hirata H, Yoshiura S, Ohtsuka T, Bessho Y, Harada T, Yoshikawa K, and Kageyama R; *Oscillatory expression of the bhlh factor hes1 regulated by a negative feedback loop.*; Science; **298**, 840; 2002.

Ho RK and Kimmel CB; *Commitment of cell fate in theearly zebrafish embryo.*; Science; **261**, 109; 1993.

Holley SA; *The genetics and embryology of zebrafish metamerism.*; Dev Dyn.; **236**, 1422; 2007.

Holley SA, Geisler R, and Nüsslein-Volhard C; *Control of her1 expression during zebrafish somitogenesis by a delta dependent oscillator and an independent wave-front activity.*; Genes Dev.; **14**, 1678; 2000.

Holley SA, Julich D, Rauch GJ, Geisler R, , and Nüsslein-Volhard C; *her1 and the notch pathway function within the oscillator mechanism that regulates zebrafish somitogenesis.*; Development; **129**, 1175; 2002.

Horikawa K, Ishimatsu K, Yoshimoto E, Kondo S, and Takeda H; *Coupling cellular oscillators: a mechanism that maintains synchrony against developmental noise in the segmentation clock.*; Nature; **719**, 441; 2006.

Huisken J and Stainier DYR; *Selective plane illumination microscopy techniques in developmental biology.*; Development; **136**, 1963; 2009.

Huisken J, Swoger J, Bene FD, Wittbrodt J, and Stelzer EHK; *Optical sectioning deep inside live embryos by selective plane illumination microscopy.*; Science; **305**, 1007; 2004.

Ishimatsu K, Horikawa K, and Takeda H; *Coupling cellular oscillators: a mechanism that maintains synchrony against developmental noise in the segmentation clock.*; Dev Dyn.; **1416**, 236; 2007.

Itoh M, Kim CH, Palardy G, Oda T, Jiang YJ, Maust D, Yeo SY, Lorick K, Wright GJ, Ariza-McNaughton L, Weissman AM, Lewis J, Chandrasekharappa SC, and Chitnis AB; *Mind bomb is a ubiquitin ligase that is essential for efficient activation of notch signaling by delta.*; Dev Cell; **4**, 67; 2003.

Jessen JR and Solnica-Krezel L; *Identification and developmental expression pattern of van gogh-like 1, a second zebrafish strabismus homologue.*; Gene Expr. Patterns.; **4**, 339; 2004.

Jessen JR, Topczewski J, Bingham S, Sepich DS, Marlow F, Chandrasekhar A, and Solnica-Krezel L; *Zebrafish trilobite identifies new roles for strabismus in gastrulation and neuronal movements.*; Nat Cell Biol; **4**, 610; 2002.

Jiang YJ, Aerne BL, Smithers L, Haddon C, Ish-Horowicz D, and Lewis J; *Notch signalling and the synchronization of the somite segmentation clock.*; Nature; **408**, 475; 2000.

Jiang YJ, Brand M, Heisenberg CP, Beuchle D, Furutani-Seiki M, Kelsh R, Warga R, Granato M, Haffter O, Hammerschmidt M, Kane DA, Mullins MC, Odenthal

J, van Eeden F, and Nüsslein-Volhard C; *Mutations affecting neurogenesis and brain morphology in the zebrafish, danio rerio.*; Development; **123**, 205; 1996.

Jülich D, Lim CH, Round J, Nicolaije C, J JS, A AD, Geisler R, Lewis J, Jiang YJ, and Holley S; *beamter/deltac and the role of notch ligands in the zebrafish somite segmentation, hindbrain neurogenesis and hypochord differentiation.*; Dev Biol.; **286**, 391; 2005.

Kageyama R, Ohtsuka T, and Kobayashi T; *The hes gene family: repressors and oscillators that orchestrate embryogenesis.*; Development; **134**, 1243; 2007.

Karsent E; *Self-organization in cell biology: a brief history.*; Nature Reviews Molecular Cell Biology; **9**, 255; 2008.

Kass M, Witkin A, and Terzopoulos D; *Snakes: Active contour models*; International Journal of Computer Vision; **1**, 321; 1988.

Keller PJ, D A, Schmidt, Wittbrodt J, and Stelzer EH; *Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy.*; Science; **322**, 1065; 2008.

Keller R, Davidson L, Edlund A, Elul T, Ezin M, Shook D, and Skoglund P; *Mechanisms of convergence and extension by cell intercalation.*; Philosophical Transactions of the Royal Society B: Biological Sciences; **355**, 897; 2000.

Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, and Schilling TF; *Stages of embryonic development of the zebrafish.*; Dev Dyn.; **203**, 253; 1995.

Kimmel CB, M R, Warga, and Schilling TF; *Origin and organization of the zebrafish fate map.*; Development; **108**, 581; 1990.

Kondo S and Miura T; *Reaction-diffusion model as a framework for understanding biological pattern formation.*; Science; **329**, 1616; 2010.

Krol A, Roellig D, Dequeant ML, Tassy O, Glynn E, Hattem G, Mushegian A, Oates AC, and Pourquié O; *Zebrafish segmentation and pair-rule patterning.*; Development; **138**, 2783; 2011.

Kruse K and Jülicher F; *Oscillations in cell biology.*; Curr Opin Cell Biol.; **17**, 20; 2005.

Kuramoto Y; *Chemical oscillations, waves, and turbulence.* (SpringerÐVerlag,, Berlin, 1984); 1st edition.

Lewis J; *Autoinhibition with transcriptional delay: A simple mechanism for the zebrafish somitogenesis oscillator.*; Curr Biol.; **13**, 1398; 2003.

Lewis J; *From signals to patterns: space, time, and mathematics in developmental biology.*; Science; **322**, 399; 2008.

Lewis J, Hanisch A, and Holder M; *Notch signaling, the segmentation clock, and the patterning of vertebrate somites.*; Journal of Biology; **8**, 44; 2009.

Li G, Liu T, Tarokh A, Nie J, Guo L, Mara A, Holley S, and Wong ST; *3d cell nuclei segmentation based on gradient flow tracking.*; BMC Cell Biol.; **8**, 40; 2007.

Lim JS; *Two-Dimensional Signal and Image Processing* (Englewood Cliffs, NJ, Prentice Hall, USA, 1990); 1st edition.

Lin G, Adiga U, Olson K, Guzowski JF, Barnes CA, and Roysam B; *A hybrid 3d watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks.*; Cytometry A; **23**, 56; 2003.

Long F, Peng H, Liu X, Kim SK, and Myers E; *A 3d digital atlas of c. elegans and its application to single-cell analyses.*; Nat Methods.; **6**, 667; 2009.

Lucy LB; *An iterative technique for the rectification of observed distributions.*; Astronomical Journal; **79**, 745; 1974.

Mara A and Holley SA; *Oscillators and the emergence of tissue organization during zebrafish somitogenesis.*; Trends Cell Biol.; **17**, 593; 2007.

Masamizu Y, Ohtsuka T, Takashima Y, Nagahara H, Takenaka Y, Yoshikawa K, Okamura H, and Kageyama R; *Real-time imaging of the somite segmentation clock: revelation of unstable oscillators in the individual presomitic mesoderm cells.*; Proc Natl Acad Sci U S A; **103**, 1313; 2006.

McInerney D, Schnell S, Baker RE, and Maini PK; *A mathematical formulation for the cell-cycle model in somitogenesis: analysis, parameter constraints and numerical solutions.*; Math Med Biol; **21**, 85; 2004.

McLachlan G and Peel D; *Finite Mixture Models.* (NJ: John Wiley and Sons, Inc.,, USA, 2000); 1st edition.

Meegama RGN and Rajapakse JC; *Nurbs snakes*; Image and Vision Computing; **21**, 551; 2003.

Megason SG; *In toto imaging of embryogenesis with confocal time-lapse microscopy.*; Methods Mol Biol.; **546**, 317; 2009.

Mehta P and Gregor T; *Approaching the molecular origins of collective dynamics in oscillating cell populations.*; Curr Opin Genet Dev.; **20**, 574; 2010.

Meijering E, Dzyubachyk O, and Smal I; *In imaging and spectroscopic analysis of living cells.*; Methods in Enzymology; **504**, 183; 2012.

Meinhardt H; *Models of Biological Pattern Formation* (Academic Press, London, 1982); 1st edition.

Meyer F; *Topographic distance and watershed lines.*; Signal Processing; **38**, 113; 1994.

Monk NA; *Oscillatory expression of hes1, p53, and nf-kappab driven by transcriptional time delays.*; Curr Biol.; **13**, 1409; 2003.

Morelli LG, Ares S, Herrgen L, Schröter C, Jülicher F, and Oates AC; *Delayed coupling theory of vertebrate segmentation.*; The HFSP Journal; **3**, 55; 2009.

Morelli LG, Uriu K, Ares S, and Oates AC; *Computational approaches to developmental patterning.*; Science; **336**, 187; 2012.

Morrison AH, Scheeler M, Dubuis J, and Gregor T; *Quantifying the bicoid morphogen gradient in living fly embryos.*; Cold Spring Harb Protoc.; **4**, 398; 2012.

Myers DC, Sepich DS, and Solnica-Krezel L; *Convergence and extension in vertebrate gastrulae: cell movements according to or in search of identity.*; Trends Genet; **18**, 447; 2002.

Niwa Y, Shimojo H, Isomura A, Gonzalez A, and Miyachi H; *Different types of oscillations in notch and fgf signaling regulate the spatiotemporal periodicity of somitogenesis.*; Genes Dev.; **25**, 1115; 2011.

Novak B and Tyson JJ; *Design principles of biochemical oscillators.*; Nature Reviews Molecular Cell Biology; **9**, 981; 2008.

Oates AC, Gorfinkiel N, Gonzalez-Gaitan M, and Heisenberg CP; *Quantitative approaches in developmental biology.*; Nat Rev Genet.; **10**, 517; 2009.

Oates AC and Ho RK; *Hairy/e(spl)-related (her) genes are central components of the segmentation oscillator and display redundancy with the delta notch signaling pathway in the formation of anterior segmental boundaries in the zebrafish.*; Development; **129**, 2929; 2002.

Oates AC, Morelli LG, and Ares S; *Patterning embryos with oscillations structure function and dynamics of the vertebrate segmentation clock.*; Development; **139**, 625; 2012.

Osher S and Sethian JA; *Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations.*; Journal of Computational Physics; **79**, 12; 1988.

Oswald A and Oates AC; *Control of endogenous gene expression timing by introns.*; Genome Biol.; **12**, 107; 2011.

Otsu N; *A threshold selection method from gray-level histograms.*; IEEE Trans. Sys., Man., Cyber.; **9**, 62; 1979.

Ozbudak EM and Lewis J; *Notch signaling synchronizes the zebrafish segmentation clock but is not needed to create somite boundaries.*; PLoS Genet.; **4**, 15; 2008.

Palmeirim I, Henrique D, Ish-Horowicz D, and Pourquié O; *Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis.*; Cell; **639**, 91; 1997.

Perona P and Malik J; *Scale-space and edge detection using anisotropic diffusion.*; IEEE Transactions on Pattern Analysis and Machine Intelligence; **12**, 629; 1990.

Peruani F, Nicola EM, and Morelli LG; *Mobility induces global synchronization of oscillators in periodic extended systems.*; New J. Phys.; **12**, 093029; 2010.

Pikovsky A, Rosenblum M, and Kurths J; *Synchronization: A universal concept in nonlinear sciences* (Cambridge University Press, UK, 2001); 1st edition.

Pourquié O; *Vertebrate somitogenesis.*; Annu Rev Cell Dev Biol; **17**, 311; 2001.

Pourquié O; *The segmentation clock: Converting embryonic time into spatial pattern.*; Science; **301**, 328; 2003.

Pourquié O; *Vertebrate segmentation: From cyclic gene networks to scoliosis.*; Cell; **145**, 650; 2011.

Press WH, Teukolsky SA, Vetterling WT, and Flannery BP; *Numerical Recipes: The Art of Scientific Computing, Section 16.1. Gaussian Mixture Models and k-Means Clustering.* (New York: Cambridge University Press,, USA, 2007); 3rd edition.

Primmett DR, Norris WE, Carlson GJ, Keynes RJ, and Stern CD; *Periodic segmental anomalies induced by heat shock in the chick embryo are associated with the cell cycle.*; Development; **105**, 119; 1989.

Raffel MC, Willert C, and Kompenhans J; *Particle Image Velocimetry: A Practical Guide* (Springer, Berlin, 1998); 3rd edition.

Reeves G, Muratov CB, Schüpbach T, and Shvartsman SY; *Quantitative models of developmental pattern formation.*; Dev Cell; **11**, 289; 2006.

Richardson WH; *Bayesian-based iterative method of image restoration.*; J. Opt. Soc. Am.,; **62**, 55; 1972.

Riedel-Kruse IH, Müller C, and Oates AC; *Synchrony dynamics during initiation, failure, and rescue of the segmentation clock.*; Science; **317**, 1911; 2007.

Roellig D, Morelli LG, Ares S, Jülicher F, and Oates AC; *Snapshot: the segmentation clock.*; Cell; **145**, 800; 2011.

Roellig D and Oates AC; *Roles for the pcp pathway during somite boundary formation.*; 2012; Ahead of Print.

Roysam B, Lin G, Abdul-Karim M, Al-Kofahi O, Al-Kofahi K, Shain W, Szarowski DH, and Turner JN; *Handbook of Biological Confocal Microscopy: Automated Three-Dimensional Image Analysis Methods for Confocal Microscopy* (SpringerScience+Business Media, New York, USA, 2006); 3rd edition.

Sbalzarini IF and Koumoutsakos P; *Feature point tracking and trajectory analysis for video imaging in cell biology.*; Journal of Structural Biology; **151**, 182; 2005.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, and Cardona A; *Fiji: an open-source platform for biological-image analysis.*; Nat Methods.; **9**, 676; 2012.

Schröter C, Herrgen L, Cardona A, Brouhard GJ, Feldman B, and Oates AC; *Dynamics of zebrafish somitogenesis.*; Dev Dyn.; **237**, 545; 2008.

Schuster HG and Wagner P; *Mutual entrainment of two limit cycle oscillators with time delayed coupling.*; Prog. Theor. Phys; **81**, 939; 1989.

Seber GAF; *Multivariate Observations.* (NJ: John Wiley and Sons, Inc.,, USA, 1984); 1st edition.

Sepich DS, Myers DC, Short R, Topczewski J, Marlow F, and Solnica-Krezel L; *Role of the zebrafish trilobite locus in gastrulation movements of convergence and extension.*; Genesis; **27**, 159; 2000.

Sezgin M and Sanku B; *Survey over image thresholding techniques and quantitative performance evaluation.*; Journal of Electronic Imaging; **13**, 146; 2004.

Shapiro LG and Stockman GC; *Computer Vision* (Prentice Hall, Englewood-Cliffs NJ, USA, 2001); 1st edition.

Solnica-Krezel L; *Conserved patterns of cell movements during vertebrate gastrulation.*; Curr Biol.; **15**, 213; 2005.

Solnica-Krezel L, Stemple DL, Mountcastle-Shah E, Rangini Z, Neuhauss SCF, Malicki J, Schier AF, Stainier DYR, Zwartkruis F, Abdelilah S, and Driever W; *Mutations affecting cell fates and cellular rearrangements during gastrulation in zebrafish.*; Development; **123**, 117; 1996.

Soroldoni D; *Tracking the oscillations of high-fidelity bac transgenes during zebrafish somitogenesis.*; Ph.D. thesis; Technischen Univesität Dresden; 2010.

Soroldoni D and Oates AC; *Live transgenic reporters of the vertebrate embryo's segmentation clock.*; Curr Opin Genet Dev.; **21(5)**, 600; 2011.

Spath H; *Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples.* (New York: Halsted Press,, USA, 1985); 1st edition.

Spemann H; *Embryonic Development and Induction* (Hafner Publishing Co Ltd, Germany, 1938); 1st edition.

Stemple DL, Solnica-Krezel L, Zwartkruis F, Neuhauss SCF, Schier AF, Malicki J, Stainier DYR, Abdelilah S, Rangini Z, Mountcastle-Shah E, and Driever W; *Mutations affecting development of the notochord in zebrafish.*; Development; **123**, 117; 1996.

Strogatz SH; *Nonlinear dynamics and chaos : with applications to physics, biology, chemistry, and engineering* (Cambridge, Mass. : Perseus Pub., UK, 1994); 1st edition.

Supatto W, McMahon A, Fraser SE, and Stathopoulos A; *Quantitative imaging of collective cell migration during drosophila gastrulation: multiphoton microscopy and computational analysis.*; Nat Protoc.; **4**, 1397; 2009.

Swinburne IA, Miguez DG, Landgraf D, and Silver PA; *Intron length increases oscillatory periods of gene expression in animal cells.*; Genes Dev.; **22**, 2342; 2008.

Swinburne IA and Silver PA; *Intron delays and transcriptional timing during development.*; Dev Cell; **14**, 324; 2008.

Takashima Y, Ohtsuka T, GonzĞlez A, Miyachi H, and Kageyama R; *Intronic delay is essential for oscillatory expression in the segmentation clock.*; Proc Natl Acad Sci U S A; **108**, 3300; 2011.

Tiedemann HB, Schneltzer E, Zeiser S, Rubio-Aliaga I, Wurst W, Beckers J, Przemeck GKH, and de Angelis MH; *Cell-based simulation of dynamic expression patterns in the presomitic mesoderm.*; J Theor Biol; **248**, 120; 2007.

Tinsley MR, Taylor AF, Huang ZY, and Showalter K; *Emergence of collective behavior in groups of excitable catalyst-loaded particles: spatiotemporal dynamical quorum sensing.*; Physics Review Letters; **102**, 158301; 2009.

Tomlin CJ and Axelrod JD; *Biology by numbers: mathematical modelling in developmental biology.*; Nat Rev Genet.; **8**, 331; 2007.

Topczewski J, Sepich DS, Myers DC, Walker C, Amores A, Lele Z, Hammerschmidt M, Postlethwait J, and Solnica-Krezel L; *The zebrafish glypican knypek controls cell polarity during gastrulation movements of convergent extension.*; Dev Cell; **1**, 251; 2001.

Turing AM; *The chemical basis of morphogenesis.*; Philosophical Transactions of the Royal Society B: Biological Sciences; **237**, 37; 1952.

Uriu K, Ares S, Oates AC, and Morelli LG; *Optimal cellular mobility for synchronization arising from the gradual recovery of intercellular interactions.*; Phys. Biol.; **9**, 036006; 2012.

Uriu K, Morishita Y, and Iwasa Y; *Traveling wave formation in vertebrate segmentation.*; J Theor Biol; **257**, 385; 2009.

Uriu K, Morishita Y, and Iwasa Y; *Random cell movement promotes synchronization of the segmentation clock.*; Proc Natl Acad Sci U S A; **107(11)**, 4979; 2010.

van Eeden F, Granato M, Schach U, Brand M, Furutani-Seiki M, Haffter P, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane D, Kelsh R, Mullins M, Odenthal J, Warga R, Allende M, Weinberg E, and Nüsslein-Volhard C; *Mu-*

tations affecting somite formation and patterning in the zebrafish, danio rerio.; Development; **123**, 153; 1996.

Vilhais-Neto GC, Maruhashi M, Smith KT, Vasseur-Cognet M, Peterson AS, Workman JL, and Pourquié O; *Rere controls retinoic acid signalling and somite bilateral symmetry.*; Nat Lett.; **463**, 953; 2010.

Waddington H; *Principles of Embryology* (Allen and Unwin, London, 1956); 1st edition.

Wallingford JB, Fraser SE, and Harland RM; *Convergent extension: The molecular control of polarized cell movement during embryonic development.*; Dev Cell; **2**, 695; 2002.

Warga RM and Kimmel CB; *Cell movements during epiboly and gastrulation in zebrafish.*; Development; **108**, 569; 1990.

Wartlick O, Mumcu P, Kicheva A, Bittig T, Seum C, Jülicher F, and González-Gaitán M; *Dynamics of dpp signaling and proliferation control.*; Science; **331**, 1154; 2011.

Webb SE and Miller AL; *Ca2+ signaling and early embryonic patterning during zebrafish development.*; Proceedings of the Australian Physiological Society; **38**, 43; 2007.

Westerfield M; *THE ZEBRAFISH BOOK; A guide for the laboratory use of zebrafish (Danio rerio)* (Eugene, University of Oregon Press,, USA, 2007); 5th edition.

Wetzel L, Morelli LG, Oates AC, Jülicher F, and Ares S; *Distributed delays in systems of coupled phase oscillators*; 2012; Ahead of Print.

Winfree AT; *The Geometry of Biological Time* (Springer-Verlag, Volume 8 series Biomathemtics, 1980); 1st edition.

Wolpert L; *Positional information and the spatial pattern of cellular differentiation.*; J Theor Biol; **25**, 1; 1969.

Wolpert L; *Positional information and pattern formation.*; Dev. Genet.; **15**, 485; 1994.

Wolpert L, Jessell T, Lawrence P, Meyerowitz E, Robertson E, and Smith J; *Principles of Development* (Oxford University Press, Oxford, 2007); 3rd edition.

Wright SJ and Wright DJ; *Selective plane illumination microscopy techniques in developmental biology.*; Methods in Cell Biology; **70**, 1; 2002.

Yamamoto A, Amacher SL, Kim SH, Geissert D, Kimmel CB, and Robertis DEM; *Zebrafish paraxial protocadherin is a downstream target of spadetail involved in morphogenesis of gastrula mesoderm.*; Development; **125**, 3389; 1998.

Yeung MKS and Strogatz SH; *Time delay in the kuramoto model of coupled oscillators.*; Physics Review Letters; **82**, 648; 1999.

Yin C, Kiskowski M, Pouille P, Farge E, and Solnica-Krezel L; *Cooperation of polarized cell intercalations drives convergence and extension of presomitic mesoderm during zebrafish gastrulation.*; Journal of Cell Biology; **180**, 221; 2000.

Yin C and Solnica-Krezel L; *Convergence and extension movements mediate the specification and fate maintenance of zebrafish slow muscle precursors.*; Dev Biol.; **304**, 141; 2007.

Acknowledgements

I would like to thank Andy, I have really learnt a lot from you and I am happy to have done my Phd under your supervision. I would like to thank my TAC members, Pavel Tomancak and Frank Jülicher for helpful comments and discussions. I would also like to specially thank Prof. Frank Jülicher for providing me an office space and computing facilities at MPI - PKS and for supporting me throughout the Phd, this thesis would not have been possible without your support. I would like to thank Hubert Scherrer-Paulus, Thomas Müller, Helmut Deggelmann from the PKS computing facility for all their technical help. I would like to thank all the people with whom I have shared my office at PKS throughout my Phd; thank you - Alireza Akbari, Susumu Shinohara, Michael Kücken, Jochen Schneider and Davoud Pouladsaz.

I would like to thank the past and present members of the Oates lab. I would like to thank Jean-Yves Tinevez who gave his valuable time for having fruitful discussions about the segmentation codes and transforming it to the open source 'Fiji'. Special thanks to Koichiro for providing the artificial data and his time and support for analyzing data from the imaging experiments. I would like to thank Guillaume and Andy for transplantation experiments, Bo-Kai for microinjections. And everyone in the lab for their insightful discussions. Many thanks to Jan Huisken for giving me the opportunity to test their SPIM facility. I would like to thank everyone at the fish facility and the Dresden fish community and the computer department at the CBG. A big thanks to the LMF facility. I would especially like to thank Jan Peychl, Dan White and Peter Gabriel Pitrone for helping me to use and learn the microscopes and for being there to solve all big and small problems related to imaging.

" I have learnt something from each one of you - thank you all."

I would like to thank all my friends and relatives all around the globe for their constant support! Last, but not the least, I would like to heartily thank my parents and my sister for their encouragement and positive thinking. I wish to thank Sebastian, for his patience and encouragement throughout the thesis. I feel extremely lucky to have you all!!

Chapter 8

Declaration

8.1 According to §5.5 of the doctoral regulations

I herewith declare that I have produced this thesis without the prohibited assistance of third parties and without making use of aids other than specified; notions taken over directly or indirectly from other sources have been identified as such. This thesis has not previously been presented in identical or similar form to any other German or foreign examination board.

The thesis work was conducted from December 2008 to May 2012 under the supervision of Prof. Dr. Frank Jülicher, Max Planck institute for Physics of Complex Systems, Dresden, Germany in the group of Dr. Andrew Oates at the Max-Planck Institute of Molecular Cell Biology and Genetics in Dresden, Germany.

I declare that I recognize the doctorate regulations of the the Fakultät für Naturwissenschaften of the Technische Universität Dresden. I declare that I have not undertaken any previous unsuccessful doctorate proceedings. Except where otherwise stated, I present my original results.

Bhavna Rajasekaran

Date Signature